AD_____

Award Number: DAMD17-93-C-3118

TITLE: Three Dimensional Structure Determination of Botulinum Neurotoxin

PRINCIPAL INVESTIGATOR: Raymond C. Stevens, Ph.D.

CONTRACTING ORGANIZATION: University of California
Berkeley, California 94720

REPORT DATE: July 1999

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1999	Final (1 Aug 93 - 30 Jun 99)

4. TITLE AND SUBTITLE
Three Dimensional Structure Determination of Botulinum Neurotoxin

5. FUNDING NUMBERS
DAMD17-93-C-3118

6. AUTHOR(S)
Raymond C. Stevens, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of California
Berkeley, California 94720

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The immediate goals of the contract on the structure and function relationship of botulinum neurotoxin are: 1) Determine the three-dimensional structure of botulinum neurotoxin at atomic resolution by x-ray crystallography. 2) Based on the structure of the neurotoxin, understand the toxins mechanism of action.

We have accomplished the first goal of determining the three-dimensional structure of the 150 kD botulinum neurotoxin serotype A. The toxin is Y-shaped, with a very long alpha-helical translocation domain forming the backbone of the structure. The translocation domain is composed almost entirely of helices, 2 of which are 95 Å in length and form a pseudo-coiled coil. The binding domain and catalytic domain are more globular in shape, located at two different ends of the translocation domain. The overall dimensions of the protein are 120 Å x 80 Å x 40 Å. A complete description of the three-dimensional structure is described in the report. Refinement and analysis of the structure are also included. To date, a total of 8 manuscripts have been published, and 3 Ph.D. thesis generated.

14. SUBJECT TERMS	15. NUMBER OF PAGES
x-ray crystallography, 3-D structure, botulinum neurotoxin	103
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500
Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18 298-102
USAPPC V1.00

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

__✓__ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

7/29/99

PI - Signature Date

Three-Dimensional Structure Determination of Botulinum Toxin
Contract: DAMD17-93-C-3118

TABLE OF CONTENTS

INTRODUCTION

Nature of Problem

To determine the 3-dimensional structure of botulinum neurotoxins and their isolated domains. Priority of the specifications are determined in consultation with the Contracting Officer's Representative (COR). The following specifications are listed in the contract section C - Statement of Work:

1) Crystallization of the 150 kD holo-botulinum neurotoxin, serotype A, B, E.

2) Determination of the 3-dimensional structure of serotype A.

3) Given the structure of serotype A, the technique of molecular replacement will be used to determine the structure of serotype B and E.

4) Crystallization of Heavy chain neurotoxin serotype A.

5) Determination of the 3-dimensional structure of the isolated heavy chain. Molecular replacement using the structure of the intact holo-neurotoxin if the structure has not changed substantially. If not, isomorphous replacement will be used.

6) Crystallization of Light chain neurotoxin serotype A.

7) Determination of the 3-dimensional structure of the isolated light chain. Molecular replacement using the structure of the intact holo-neurotoxin if the structure has not changed substantially. If not, isomorphous replacement will be used.

8) Similar studies on other serotypes.

Based on the consultation with the COR, priority is to the 3-dimensional crystal structure of the 150 kD holo-botulinum neurotoxin serotype A.

Background of Previous Work

Before funding of the contracted work for the structure determination of botulinum neurotoxin serotype A, we determined crystallization conditions of botulinum neurotoxin serotype A (Stevens et al, 1991) and found the most suitable crystal form of serotype A for X-ray analysis are bipyramidal shaped crystals that crystallize in the hexagonal space group $P3_121$ (or $P3_221$) with one monomer per asymmetric unit. The unit cell dimensions are a = b = 170.5 Å, c = 161.7 Å. Native and derivative data sets have been collected in house (Room 405 Stanley Hall) at the University of California, Berkeley using a shared (with 3 other research groups) Rigaku RU-200 rotating anode generator and R-axis Image Plate area detector system.

During the first 2 years of the 3 year contract, we reported that we had worked out the conditions to further purify the protein (isoelectric focusing and aggregation), stabilize the protein (zinc acetate, proteolytic inhibitors, ganglioside), and improve the crystal quality of the neurotoxin protein. We very carefully worked out conditions to collect complete, accurate, and non-deteriorating x-ray diffraction data use flash freezing techniques to freeze the protein crystals at -170°C during x-ray irradiation. We

collected a large amount of both native (7 sets) and heavy atom derivative data (20 sets) that is used to phase the x-ray diffraction pattern to yield the 3-dimensional structure of the neurotoxin. Data have been collected at the Brookhaven NSLS and Stanford SSRL synchrotron radiation facilities. As a feasibility study, we initiated crystallization trials of serotype B and E and we have obtained single crystals (completion of Contract Goal C2).

Purpose of Present Work

To complete the contracted work on the 3-dimensional structure determination of botulinum neurotoxin and understand the toxins structure, function, and mechanism of action. Specifically we will locate the heavy atom positions of data collected and phase the diffraction pattern to allow us to trace the electron density of the protein molecule.

Methods of Approach

Using the technique of multiple isomorphous replacement, the 3-dimensional structure of botulinum neurotoxin will be solved using diffraction data from native and heavy atom derivative data sets that are collected in house or at synchrotron facilities. The technique requires one to:

1) Bind heavy atoms to derivative the protein crystal and locate the position of the heavy atom using Patterson methods. The derivatized protein crystals must be isomorphous with the native protein crystals (except for the heavy atom itself) for the information to be interpretable.

2) Once the position of the heavy atoms have been located, approximate phases can be calculated to allow one to observe the electron density of the protein molecule. The program MLPHARE from the CCP4 program suite (CCP4, 1979, The SERC (UK) Collaborative Computing Project No. 4, a Suite of Programs for Protein Crystallography, distributed from Daresbury Laboratory, Warrington, WA4 4AD, UK) will be used to combine and phase heavy atom derivative data. All computations will be conducted on a Silicon Graphics OCTANE graphics workstation.

3) Fitting of electron density will be made on a Silicon Graphics OCTANE graphics workstation using the program O (version 6.1; Alwyn Jones, 1997).

4) Upon completion of fitting the electron density, the experimental model will be refined using the program CNS (version 0.5c; Brunger, 1997) installed on a Silicon Graphics OCTANE graphics workstation.

An alternative to isomorphous replacement is the method of molecular replacement which depends on the presence of related structures in different crystals. Proteins which are homologous and have closely similar structures are particularly useful. The near identity of the structures implies relations between different structure amplitudes and phases which are helpful in solving phase problems. Serotypes A, B and E are similar as well as dissimilar (pharmacologically similar, antigenically different). Hence, once a structure is obtained for one serotype, the analysis of crystals of other serotypes can be aided by the molecular replacement technique. This is based on the assumption that the overall tertiary structure of all three serotypes are similar. If this is not the case, then a search for heavy atom derivatives will have to be conducted for all three serotypes.

BODY - Experimental Methods

Although botulinum neurotoxin serotype A crystallizes quite easily, the ability to collect x-ray diffraction data on protein crystals that are all "locked" into the same structure inside the crystal has become a hurdle in our efforts to determine the 3-dimensional structure. In layman's terms, botulinum neurotoxin molecules have the ability to conform to several different conformational structures caused by "floppy" or disordered regions of protein structure. This problem makes it difficult to merge together native data sets as well as heavy atom derivative data sets. This problem makes it difficult to use the techniques of isomorphous replacement with heavy atom derivatives since one data set is not similar enough to the next. To avoid this problem we have taken several different measures that are described below. These approaches have proven successful to other researchers that have encountered this problem. The problem of non-isomorphism is not uncommon in crystal structure determinations, particularly for neurotoxins (i.e. crystal structure determinations for cholera toxin, diphtheria toxin, pertussis toxin).

1) Neurotoxin Data Collection (Table I).

Data have been collected on approximately 150 data sets in house with a series of different transition metals, sugar molecules, gangliosides, inhibitors to zinc proteases, different pH conditions, and various salts to resolve the issue of the protein molecules packing in different orientations in the crystal lattice. Each different compound is used to collect data on two different crystals in order to confirm the effects that are observed. The addition of additives is to "lock" the protein molecules into one conformation. This will allow the technique of MIR to be used for the crystal structure determination. Perhaps the most critical step in determining the crystal structure was the change from the crystal form that was being used in previous years (~1.1 M sodium formate, 100 mM Hepes pH 7) to a new crystallization conditions (200 mM Mg acetate and ~10% PEG4K, pH 7). The toxin has been crystallized in 5 different crystal morphologies, all five belonging to the same space group P3121. The common theme of acetate in the buffer and derivative most likely helped with isomorphism.

2) Heavy Atom Derivative Data Collection (Table I).

In addition to the numerous data sets collected in house, we have also collected data at two different synchrotron sources and tuned the wavelength to the absorption edge of the metal that we have bound to the protein. This approach allows us to determine the 3-D structure of the neurotoxin using anomalous dispersion of the heavy atom and maximizing the heavy atom signal. The additional data obtained from the anomalous signal can be enough to determine the crystal structure with a single derivative or at least increase the amount of usable data from one derivative. Of particular use in our study, are acetate based derivatives, since the crystal form finally used in the structure determination was from PEG-400 and Mg acetate mother liquor. Based on prior experiments, 5 free thiol groups are accessible and therefore mercury and gold derivatives are good choices for heavy atom derivatives of the toxin. The samarium and uranium derivatives bind to positively charged pockets. Merging statistics in Table I include all data to the highest resolution bin without throwing away any data.

Table I - Heavy Atom X-ray Data collected on Botulinum Neurotoxin serotype A (Crystal space group P3121)

DataSet Name	Resolution Å	Unit Cell (a,c) Å	#Refl. collected	#Refl. unique	Complete %		I/sig %	R_{merge} %	$R_{mergetonative}$ %
Synchrotron Radiation Data Collection Trips									
(SSRL March 97)									
Native	3.6	170.8, 161.1	97,575	28,963	91.8 (76.4)	9.0	9.1 (37.3)		-
HgCl	4.5	169.8, 160.9	87,457	15,830	97.2 (96.0)	6.5	12.9 (20.1)		21.7
HgAc	4.2	169.9, 161.2	86,260	19,650	97.8 (94.3)	7.1	9.1 (24.5)		21.0
SmAc	4.5	170.6, 161.0	22,797	12,516	78.5 (78.2)	7.9	7.6 (17.4)		14.9
UrAc	4.5	170.7, 161.1	22,795	11,796	71.9 (73.6)	7.8	9.3 (16.1)		15.2
KAuCl	6.0	170.7, 161.0	19,894	6,080	93.7 (95.5)	10.7	9.6 (26.2)		15.2

Highest resolution statistics are shown in parentheses

Abbrevations for heavy atom derivatives - all soaks were completed at room temperature. All data collected at 4°C.

HgCl	3 mM soak for 6 hours with methyl mercury chloride
HgAc	1mM (saturated) for 6 hours with mercury acetate
SmAc	50 mM for 12 hours with samarium acetate
UrAc	0.5 mM for 6-10 hours with uranyl acetate
KauCl	1 mM for 8-12 hours with potassium gold tetrachloride

3) Location of Heavy Atom Sites for Phasing

Interpretation of Patterson maps to locate the positions of the heavy atoms has been completed, the following heavy atom sites along with the derivative name are listed below. The derivative abbreviations are listed in Table I.

Metal	x	y	z	Occ.	B-factor
HgAc	0.9402	0.8193	0.0162	0.77	20.0
HgAc	0.6051	0.3112	0.0811	0.46	20.0
HgAc	0.0852	0.7199	0.3591	0.48	20.0
HgAc	0.1275	0.0563	0.6916	0.40	20.0
HgAc	0.3610	0.3171	0.3135	0.50	20.0
HgAc	0.2829	0.7060	0.2922	0.42	20.0
SmAc	0.4911	0.8897	0.0603	0.76	20.0
SmAc	0.4374	0.2335	0.4240	0.70	20.0
SmAc	0.0881	0.8896	0.1364	0.43	20.0
UAc	0.4952	0.8924	0.0593	0.55	20.0
UAc	0.4397	0.2367	0.4331	0.38	20.0
UAc	0.4095	0.2901	0.2264	0.35	20.0
AuCl	0.7206	0.4196	0.0368	0.93	20.0
HgCl	0.941	0.820	0.016	0.81	20.0
HgCl	0.364	0.280	0.308	0.64	20.0
HgCl	0.128	0.056	0.694	0.30	20.0
HgCl	0.690	0.294	0.412	0.43	20.0
HgCl	0.954	0.641	0.982	0.54	20.0
HgCl	0.424	0.717	0.963	0.51	20.0

4) Phasing of Diffraction Data from MLPHARE (CCP4 Program Suite).

Data were phased with 5 derivatives to 5.0 Å resolution, and the phases were extended to 3.6 Å resolution with the program DM (CCP4 program suite) using the solvent flattening/histogram matching algorithm. Although the mercury derivatives and the gold derivatives have sites in common, the occupancy differences appears to allow them to complement one another, rather than interfere with one another in the phasing statistics.

<4SSQ/LL>	Resol	Nref_a	DISO_a	LOC_a	PhP_a	CullR_a	Nref_c	DISO_c	LOC_c	PhP_c	CullR_c
0.009	10.75	634	252.0	178.0	1.19	0.71	134	350.9	248.6	0.79	0.71
0.014	8.37	1142	229.7	154.5	1.30	0.67	190	325.1	210.7	0.93	0.65
0.021	6.86	1831	181.7	112.4	1.73	0.62	229	243.9	163.3	1.10	0.67
0.030	5.81	2685	153.0	96.5	1.96	0.63	278	228.8	158.8	1.16	0.69
0.039	5.03	1173	160.0	110.3	1.69	0.69	105	202.3	152.2	1.32	0.75
0.051	4.44	0	0.0	0.0	0.00	0.00	0	0.0	0.0	0.00	0.00
0.063	3.98	0	0.0	0.0	0.00	0.00	0	0.0	0.0	0.00	0.00
0.077	3.60	0	0.0	0.0	0.00	0.00	0	0.0	0.0	0.00	0.00
TOTAL		7465	181.3	118.4	1.64	0.65	936	266.6	182.6	1.03	0.68

5) Determination of molecular envelope and protein boundary

The solvent content is estimated at 67% assuming one molecule per asymmetric unit (6 molecules per unit cell). In the previous annual report, the molecular boundary and packing arrangement were illustrated. Based on these results, a three-dimensional model could be built into the electron density.

6) Fitting Atomic Model into Experimental Electron Density

The steps in the model building are as follows:

a) Place poly-alanine secondary elements into electron density, starting with the alpha-helices, and then beta-strands.

b) Determine directionality of the secondary elements (N-terminus versus C-terminus).

c) Connect secondary elements where possible.

d) Assign sequence residue to poly-alanine trace for large side chains (Trp, Tyr, Phe).

e) Complete as much of the trace and sequence assignment as possible.

f) Refine the initial model into the electron density.

g) Perform successive rounds of model building and refinement, until the structure has converged to an acceptable crystallographic R-value and free R-value.

7) Increased resolution structure, interpretation of the Botulinum Neurotoxin 3-D structure, and sequence homology analysis (Appendix Publications 5 and 7)

Sequence alignments and superposition on the BoNT/A structure: Nine CNT sequences were aligned to that of BoNT/A using CLUSTAL W (Thompson *et al.*, 1994). A gap opening penalty of 10, gap extension penalty of 0.05, and gap separation distance of 8 were used with the BLOSUM62 matrix (Henikoff & Henikoff, 1996). This matrix was also used in the display in Figure 3 using a global score of 0.15. The Risler matrix was used to generate per residue scores for the alignment (Risler *et al.*, 1988). These scores were output in the B-factor column of the BoNT/A coordinates facilitating the display of sequence conservation on the three-dimensional structure. Residues with scores 68 and higher were colored purple while those below 68 were colored yellow. The number 68 was empirically determined to best match the display of sequence conservation.

Toxin Structure at -170 °C: The protein was concentrated to 5 mg/mL in 10 mM HEPES, pH 7 and 100 mM KCl. This stock was mixed 1:1 with a mother liquor of 30% PEG 600 and 100 mM Tris pH 7 and allowed to crystallize by hanging drop vapor diffusion. The crystal formed overnight and grew to a size of about 100 x 200 x 520 μm over the course of a week. The crystal was washed briefly in a solution made from 2 μL mother liquor and 2 μL 30% PEG 600 and plunged into liquid nitrogen. Data were collected at SSRL beamline 9-1 and processed with DENZO yielding data 99.4 % complete to 3.3 Å with 98.5% completeness in the last bin. The data were 5.8-fold redundant with an Rmerge of 8.5% overall. After orienting the molecule with AMoRe, the structure was readily refined into the frozen data using CNS, yielding a final Rwork of 26% and Rfree of 31%. While the processing and refinement statistics indicate that the quality is lower than the structure solved at 4°C, this data did bring in density for residues 561-568.

Overall Structure

The structure of BoNT/A was recently solved to 3.3 Å resolution using data collected from multiple crystals at 4°C (Lacy *et al.*, 1998). The structure (Figure 1B) showed that the binding domain was structurally similar to the TeNT binding domain (Umland *et al.*, 1997), and could be divided into two subdomains, an N-terminal β-barrel and a C-terminal β-trefoil fold. The translocation domain fold is markedly different from the folds observed in other toxins that undergo pore formation and translocation (Lacy & Stevens, 1998). Most notable are a kinked pair of α-helices, 105 Å in length, and a 54 residue 'belt' which wraps around the perimeter of the catalytic domain. The translocation domain occludes access to a large negatively charged cleft leading into the active site zinc of the catalytic domain. The zinc appears to be directly coordinated by His 222, His 226, Glu 261 and a water-mediated coordination through Glu 223.

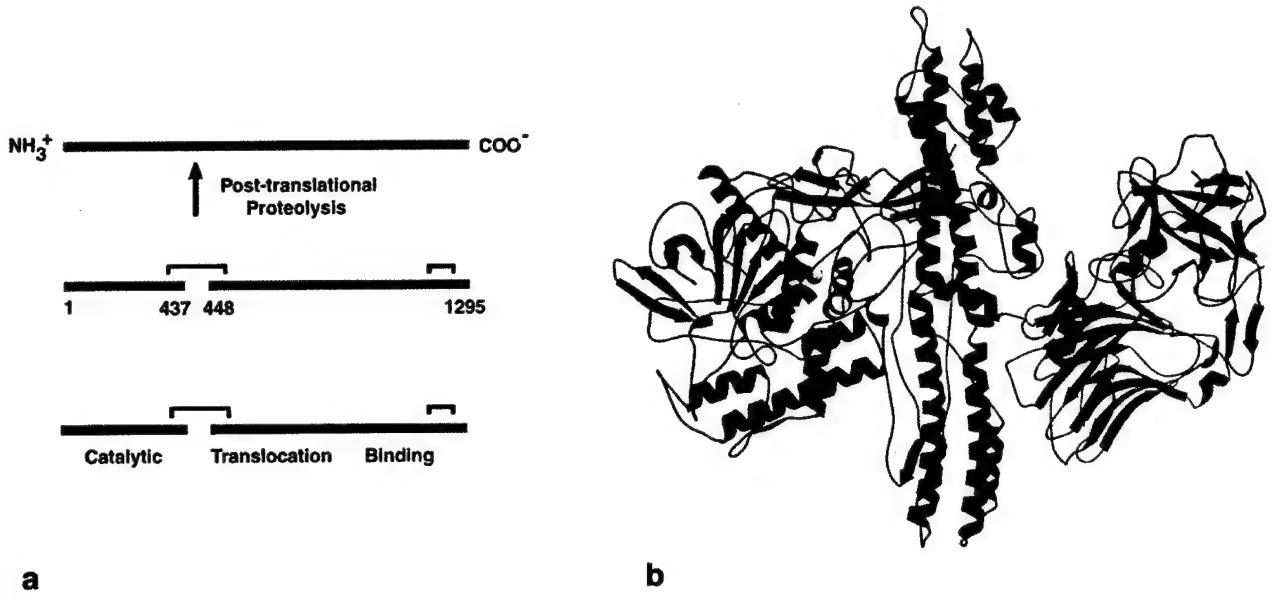


Figure 1: a) The 150 kDa toxin is post-translationally proteolyzed to form the activated di-chain molecule (arrow in top figure illustrates where proteolysis occurs). Two disulfides exist, one that links the two chains and the other in the C-terminal half of the binding domain. The three functional domains are each ~50 kDa and correspond to the catalytic domain (1-437), the translocation domain (448-872), and the binding domain (873-1295) where the numbers refer to the BoNT/A sequence. b) A backbone trace of the BoNT/A structure with the catalytic domain in purple, the translocation domain in green, the N-terminal binding sub-domain in pink, and the C-terminal binding sub-domain in blue. The catalytic zinc is represented as a sphere with the helix containing the HEXXH zinc binding motif in red. Helix α22 of the C-terminal binding sub-domain is colored orange to show the relative orientation of the putative ganglioside binding site to the rest of the molecule. The structure represents residues 1-431 and 450-1295 with residues 432-437 and residues 448-449 at the site of proteolytic cleavage being disordered. This figure was generated using MOLSCRIPT and Raster3D (Kraulis, 1991; Merritt & Bacon, 1997).

The CNT sequence divergence is substantial when one considers that the family of proteins has maintained striking similarities in a complex series of functions. In spite of the divergence though, it is reasonable to presume that the BoNT/A structure will be representative for all of the CNT's. Firstly, there is a greater conservation in predicted secondary structure than in the primary structure (Lebeda & Olson, 1994). Secondly, the structure of the binding domain overlays with the tetanus neurotoxin binding domain with a root mean square deviation (rmsd) of 1.5 Å for 363 C-alpha atoms. Lastly, the functional requirements, especially those of acid-triggered pore formation and translocation are rigorous and are not likely to tolerate significant structural changes. It is tempting to think that viewing the primary structure differences within the tertiary structure will lend insight into the molecular determinants of differences between the CNT's. The key differences seem to be an ability to discriminate between protein receptor molecules and in the enzyme specificity for one of eight cleavage sites in the three SNARE proteins. An analysis of the sequence conservation can also help identify regions required in the common functions of ganglioside binding, pore formation, translocation, and catalysis.

The binding domain

The structure of the TeNT binding domain has been solved by independent investigators to 2.7 Å (Umland et al., 1997) and 1.5 Å (PDB accession code 1A8D). The main chains of the structures differ

only in the orientation of two surface loops, located in crystal packing interfaces. As both structures were obtained in the same space group, these packing interface differences, along with the temperature and x-ray source differences are likely to account for the differences in resolution. The structure of BoNT/A at 4°C, and now –170°C, reveals the orientation of the binding domain and the accessibility of the surface loops in the presence of the translocation domain (in comparison to the TeNT binding domain only structure). The binding domain tilts away from the plane of the catalytic and translocation domains by ~60°. It projects away from the long axis of the translocation domain by ~40° in the structure from 4°C data and by ~45° in the structure refined into frozen data. This observation is not surprising given the relatively small interface between the binding and translocation domains. (The binding domain buries ~400 $Å^2$ of the translocation domain while the translocation domain buries ~480 $Å^2$ of the binding domain.) The relative orientation of this domain could vary even more substantially under physiological conditions. The two sub-domains of the binding domain are linked by an α-helix (α21) and create a cleft in their interface (Figure 2A, 2C). The C-terminal subdomain buries ~500 $Å^2$ of the N-terminal subdomain while the N-terminal subdomain buries ~540 $Å^2$ of the C-terminal domain. The contacts between the two subdomains are made through loops such that this interface could also be flexible, although the subdomain β-strands align almost identically in the three structures.

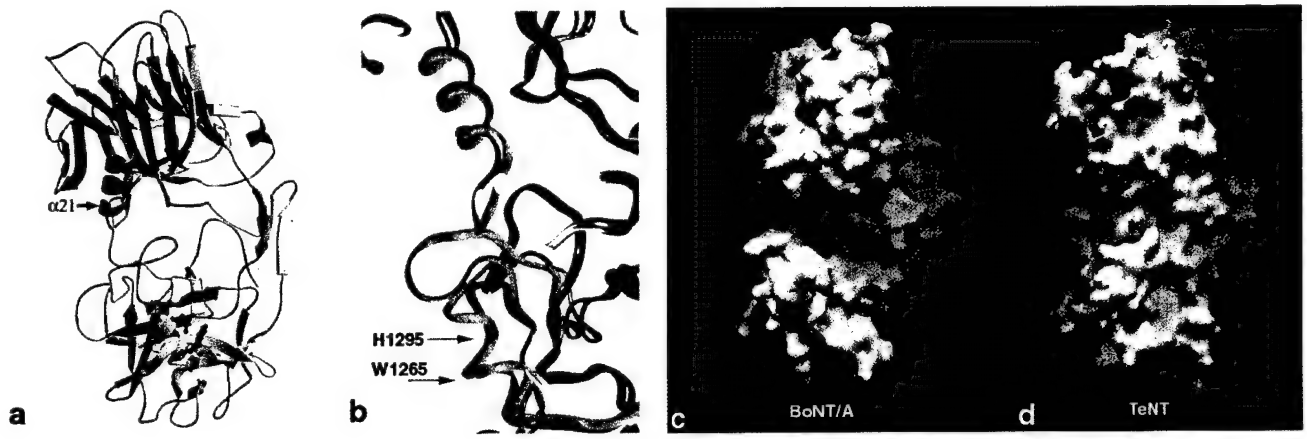


Figure 2: a) The conserved residues of the BoNT/A binding domain are colored purple indicating that the N-terminal sub-domain (top) is more highly conserved than the C-terminal sub-domain (bottom). The second disulfide in BoNT/A, green, is located behind α22. This helix, with residues W1265 and Q1269 colored in red, is thought to form part of the ganglioside binding site. b) A superposition of the BoNT/A (yellow) and TeNT (red) binding domains at the ganglioside binding site and a loop difference between the two structures. c) The two subdomains of the BoNT/A binding domain are connected by α21 and form an electrostatically positive cleft in their interface. d) The cleft in TeNT is relatively shallow and neutral, largely due to the loop difference shown in panel (b).

The alignment of BoNT/A with TeNT and the six other BoNT serotypes from the three identified organisms is shown in Appendix publication 8 along with the secondary structure assignments and solvent accessibility of residues in the BoNT/A structure (Lacy & Stevens, 1999). There is significantly greater sequence homology within the N-terminal binding subdomain as compared to the C-terminal subdomain (Figures 2A). The majority of these highly conserved residues in the N-terminal subdomain however point inward, seeming to preserve the hydrophobic core of the β-barrel. As most of the antigenicity conferring serotype uniqueness seems to arise in the binding domain (Chen *et al.*, 1997), it is not surprising that its surface residues vary dramatically among the CNT's.

While a protein receptor binding site is still under investigation, the ganglioside binding site is better understood. A binding study using trypsin digests in BoNT/A indicates that the 30 C-terminal residues of the C-terminal sub-domain are involved in binding the GT1b ganglioside (Shone *et al.*, 1985). A photoaffinity labeling study using a novel ganglioside probe implicates this same region in TeNT, showing labeling of H1295 (Shapiro *et al.*, 1997). The functional significance of this region has also been identified in BoNT/E (Kubota *et al.*, 1997). A monoclonal antibody capable of neutralizing BoNT/E in mice also bound the peptide YLTHMRD. Comparison of this sequence location in BoNT/E aligns to residues 1292-1298 of TeNT and residues 1266-1272 of BoNT/A. Most recently, a binding assay that followed tryptophan fluorescence in BoNT/A showed fluorescence quenching upon binding the ganglioside (Kamata *et al.*, 1997). However, with three tryptophans within the 30 C-terminal residues it was impossible to narrow the binding site further. The BoNT/A structure shows that of the three tryptophans, only one (W1265) is solvent accessible. This highly hydrophobic residue is fully exposed to solvent and makes contact with Q1269, the residue in BoNT/A that aligns to H1295 of TeNT, implicated in the photoaffinity labeling study. These two residues are located in the i and i+4 positions of helix 22, beneath a loop whose orientation varies dramatically between BoNT/A and TeNT (Figure 1B and 2B). While in BoNT/A this loop points out, creating a deep positively charged cleft between sub-domains, the loop in TeNT folds in, creating a shallow, more neutral cleft. This surface difference could have a role in the alternate localization of the two toxins.

The translocation domain

The translocation domain is able to form channels in artificial bilayers (Blaustein *et al.*, 1987; Donovan & Middlebrook, 1986; Hoch *et al.*, 1985) and in cell membranes (Sheridan, 1998). Visualization of these channels using electron cryomicroscopy suggests that the channels may be formed by the oligomerization of BoNT to form a tetramer (Schmid *et al.*, 1993). Efforts to identify the pore forming segment(s) of the CNT's have focussed on identifying amphipathic sequences capable of spanning the membrane (Lebeda & Olson, 1995; Montal *et al.*, 1992). Three such sequences were identified using the MOMENT algorithm for hydrophobic moments (595-614, 625-647, and 648-691) (Lebeda & Olson, 1995). A peptide representing part of one of these sequences 659-681, was shown to form channels in planar lipid bilayers (Oblatt-Montal *et al.*, 1995). This sequence could possibly oligomerize to form a four-helix-bundle in the membrane. The structure of BoNT/A, solved at pH 7, does not refute or support this hypothesis. The previously identified amphipathic sequences do not correspond to the long pairs of kinked α-helices observed in the translocation domain. Instead, the sequences preceed these helices and adopt primarily extended loop conformations (Figure 3A).

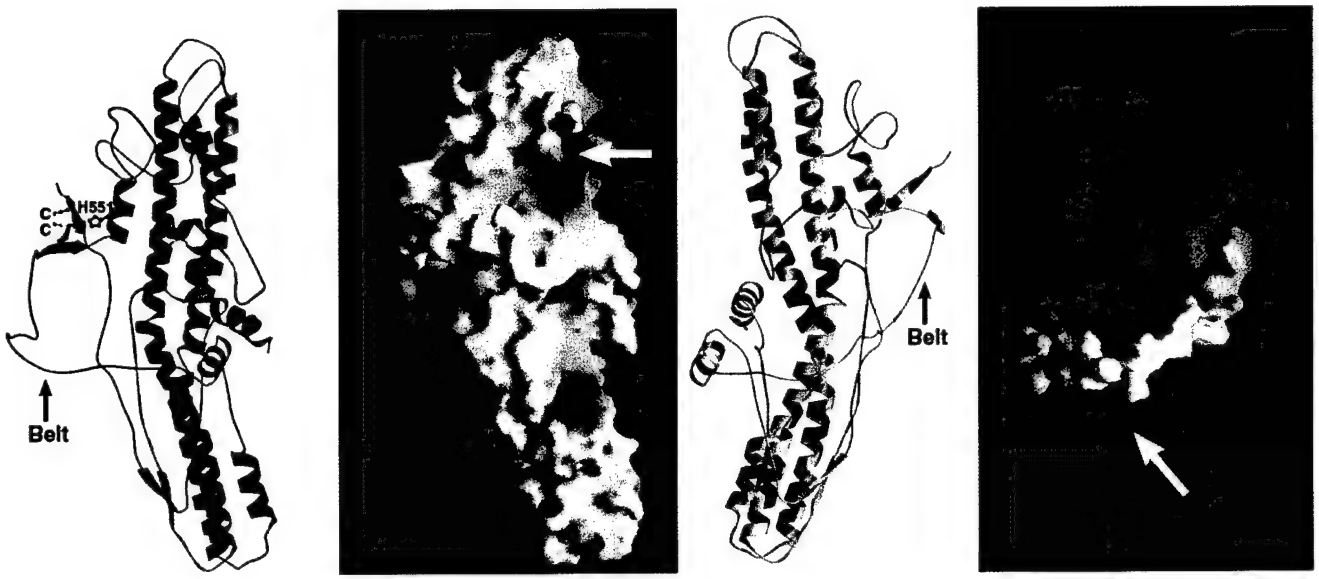


Figure 3: Panels (a) and (b) show identical orientations of the BoNT/A translocation domain and are related to panels (c) and (d) by a rotation of ~180 around the long axis of the helices. In panels (c) and (d) the translocation domain 'belt' is protruding out of the plane of the page. In panel (a) H551 is shown in purple pointing toward the disulfide (yellow) linking the translocation (green) and catalytic (shown partially in blue) domains. The loop which was disordered in the 4 °C data is shown in black. The three amphipathic sequences possibly involved in pore formation are colored in orange and red, with conserved residues colored in red. These sequences form part of a conserved electrostatic cluster shown with an arrow in panel (b). Panel (c) shows the overall sequence conservation in the translocation domain structure. The conserved residues, shown in purple, largely act to hold the helical arrangement together, the exception being the conserved residues in the extended loops of the amphipathic sequences, shown in panel (a). The arrow in panel (d) points to a cluster of charges conserved at the interface between the main body of the translocation domain and the beginning of the translocation domain 'belt'. Panels (a) and (c) were made using MOLSCRIPT (Kraulis, 1991), while the electrostatic potential surfaces of panels (b) and (d) were made using GRASP (Nicholls *et al.*, 1991).

The conservation among the CNT's in the translocation domain is shown structurally in Figure 3C. While virtually no sequence conservation is observed in the region of the belt, the sequence conservation is fairly high throughout the rest of the translocation domain. The majority of the hydrophobic and hydrophillic conserved residues are evenly spaced along the long helices (α14, α15, α16, and α17). The organization seems to preserve the packing between each other and between the smaller helices at either end of the domain (α12 and α18). The only area outside of these helices of particular note is the region from 622-644. This sequence is largely loop, with one 3-10 helix (η10) and is part of an amphipathic sequence, 625-647, previously identified (Lebeda & Olson, 1995). The loop extends away from the domain pointing toward the binding domain but is not involved in any inter-domain contacts. To our knowledge, this sequence has not been investigated for its ability to form pores in membranes.

The critical follow-up question after identifying the pore-forming segment is to identify the molecular mechanism by which pH triggers this sequence to change structure and form a membrane-spanning channel. The translocation domain is geared to sense this environmental variable, as the calculated pI for its sequence is 4.66 (as compared to 9.1 for the binding domain and 6.0 for the whole toxin). This low value is attributable to the first half of the domain containing the amphipathic loops (548-685: pI=4.6) as opposed to the second half containing the kinked pair of α-helices (686-872: pI=8.5). The three-dimensional charge distribution indicates several clusters of negative charges on both sides of the surface (Figure 3B and 3D). One cluster (Figure 3B) is comprised of D612, D615, E616,

E619, and E588. Of these residues D612 is strictly conserved, E616 and E588 are largely conserved and E619 is consistently polar. The clustering of negative charge is expected to raise the pKa of these residues, making them titratable at endosomal pH (4.5-5.5). We therefore propose that these residues and possibly any histidines will effect the structural changes required for pore formation.

The translocation domain only contains two histidines (H551 and H560) located in and after α10 respectively such that they immediately follow the translocation domain belt (492-545) and preceed a disordered loop in the structure. While H560 is solvent accessible, H551 is buried at the interface of the catalytic and translocation domains within 7 Å of their connecting disulfide. The idea that H551 protonation could disrupt this interface is supported by the fact that the structure to follow is likely to be very flexible. The original data for BoNT/A showed high B-factors for residues 560-585, with 561-568 completely disordered. The newer data, collected from a single frozen crystal, revealed density in this area although the B-factors for residues placed in this density remain high (Figure 3A). This region likely represents an area of inherent flexibility and is proximal both to the negatively charged cluster and the buried histidine. This feature could facilitate the structural changes that accompany pore formation as well as the translocation of the unfolded catalytic domain through the pore. The role that the 'belt' might play in translocation is unknown, although with movement in this loop region it is not implausible to consider that the 'belt' could move as well. A second electrostatically charged cluster on the other side of the translocation domain (Figure 3D) might also contribute to 'belt' movement. This cluster is formed by both positive and negative residues of α14, (K700, E703, K704, D706, E707, and K710), along with negative charges at the beginning of the belt (D473, E478, E487, and E490). Protonation at this interface could interrupt interactions holding the 'belt' to the main body of the translocation domain allowing for the translocation and release of the catalytic domain. While the domains are tethered by a disulfide and the interface between the catalytic and translocation domains (not including the belt) includes ~1,190 $Å^2$ of buried surface area, the non-covalent contacts between the belt and catalytic domain are not substantial. The 54 residue belt buries an additional ~1620 $Å^2$ of the catalytic domain surface area. This translates to ~30 $Å^2$ per residue of the belt, considerably lower than the ~200 $Å^2$ buried by an isoleucine in a well-packed protein core. The energy required for removing the belt is probably even lower than that indicated by the buried surface area since the belt is presumably unstructured in the absence of the catalytic domain, thus adding an entropic cost to their association.

The catalytic domain

Previous alignments have indicated that the catalytic domains share up to 36.5% sequence identity, the exception being the catalytic domains of BoNT/B and TeNT which share 51.6% sequence identity (Kurazono *et al.*, 1992). All sequences contain a HEXXH sequence motif, typical of many zinc proteases. The catalytic domain sequence alignment also shows the strict conservation of two Asp-Pro bonds, one immediately preceding the HEXXH sequence (D215 and P216) and the other at the N-terminus of the molecule (D11 and P12). A third Asp-Pro bond is present in all of the sequences, except BoNT/A where S74 exists instead of a proline. Asp-Pro bonds from BoNT/A have been shown to hydrolyze at pH 5 (DasGupta & Evenson, 1992; DasGupta & Tepp, 1991), a fact that may bear relevance given the low pH environment of the endosome and the need to translocate a 50 kD protein. However, Kurazono and co-workers noted that a chain truncation at position 10 inactivates the toxin and therefore this requires further investigation.

Other than the zinc protease HEXXH sequence motif, the catalytic domains share no sequence similarities to proteins outside of the CNT family. While a DALI search showed that thermolysin, a well-characterized zinc protease, bore the highest structural identity to the BoNT/A catalytic domain, this similarity was weak with a Z-score of 4.6 for 139 residues (Lacy et al., 1998). Nevertheless, a visual comparison of the two enzymes (Figure 4) is useful because like BoNT/A, thermolysin also coordinates

the active site zinc with two histidines (His 142 and His 146), a glutamate (Glu 166), and a water-mediated glutamate (Glu 143) (Colman *et al.*, 1972). Structural similarities include the helix containing the HEXXH sequence (α3) and a four-stranded β-sheet (β3, β6, β7, and β8). This accounts for the structural presentation of the two histidine ligands and the glutamate which coordinates the activated water in catalysis. The presentation of the fourth ligand is intriguing as thermolysin presents the Glu 166 ligand with a single helix, while BoNT/A has generated a similar presentation by folding with two smaller helices (α4 and α8) end-to-end. As the two BoNT/A helices point in opposite directions it is uncertain if maintaining this helical periodicity plays a role in the Glu 261 presentation. More interestingly, it supports this activity as a likely example of convergent evolution.

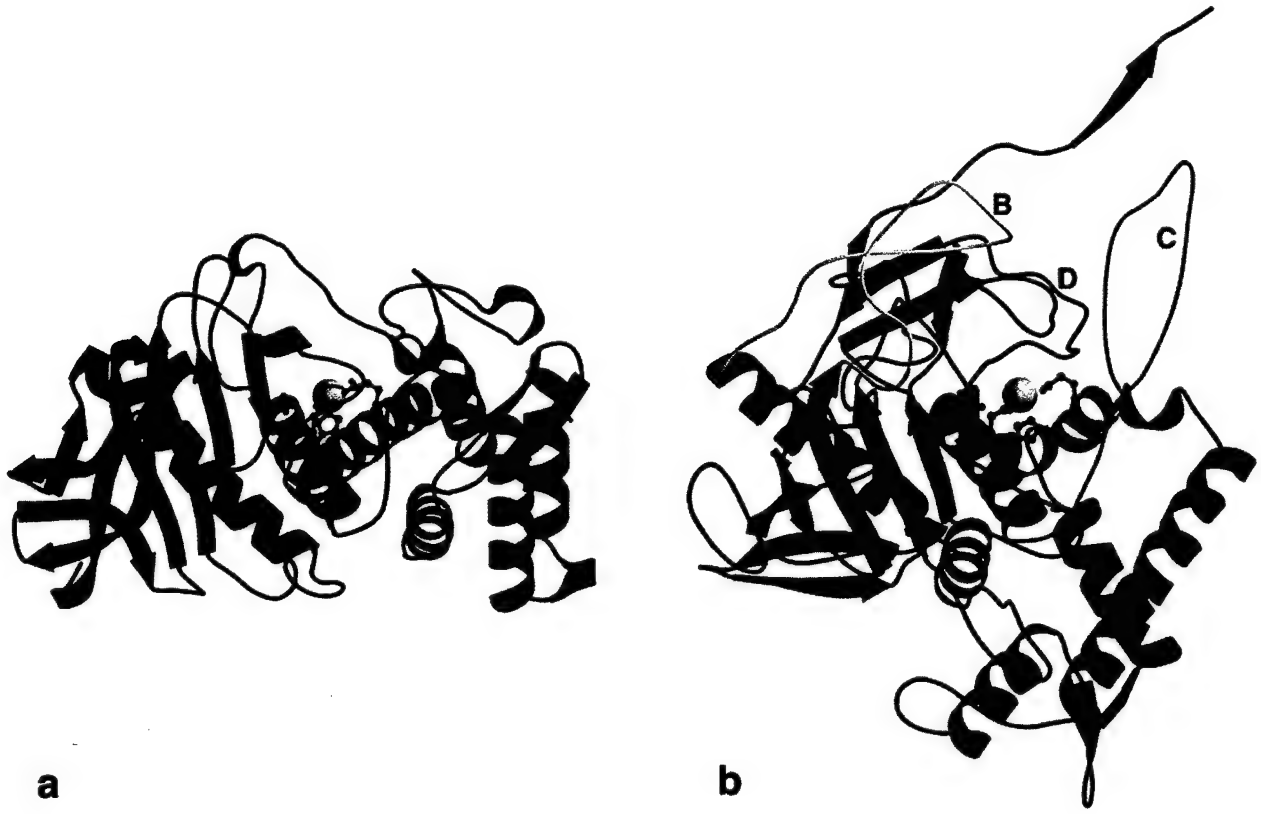


Figure 4: A comparison of thermolysin (a) and the BoNT/A catalytic domain (b). The helix containing the HEXXH zinc protease motif is colored red while the helix presenting the second Glu is colored in gold. The topology of the secondary structural elements differs such that in BoNT/A, the shorter green and gold helices come together to form the structural equivalent of the single gold helix in thermolysin. Loops B, C, and D of BoNT/A are colored in yellow, pink, and orchid, respectively, and form boundaries of accessibility to the active site zinc.

The fold comparison highlights the likely cleft by which the substrate gains accessibility to the active site. This channel of accessibility is bordered by three loops in BoNT/A designated B (47-80), C (231-259), and D (356-371) (Figure 4b). The loops in BoNT/A are longer than those in thermolysin are and result in a more buried active site, although these loops could be flexible or oriented differently in the absence of the translocation domain. A second avenue of active site access would be for substrate to enter from a channel seen arially both in Figure 4b and Figure 5a. The presence of the translocation

domain belt and the main body of the translocation domain (Figure 5b) occlude both channels of accessibility in the holotoxin.

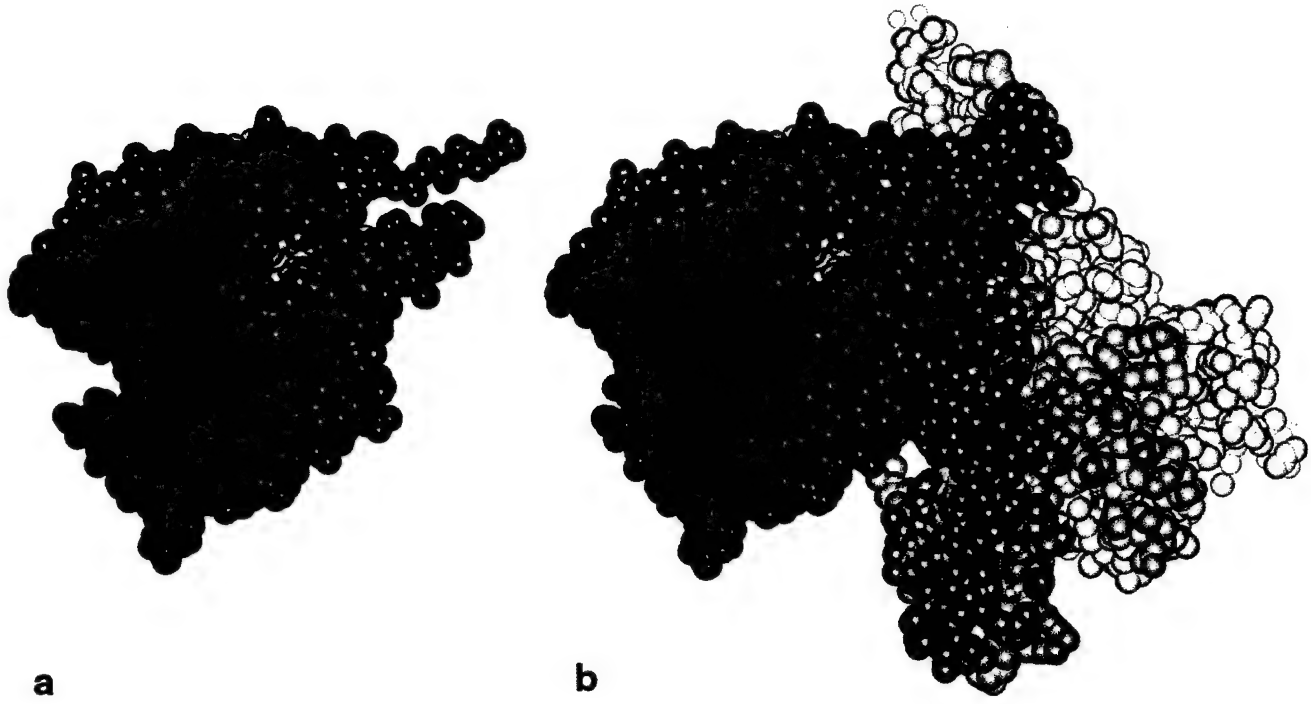


Figure 5: A CPK model of the BoNT/A catalytic domain (a) and BoNT/A holotoxin (b) showing how accessibility to the buried active site zinc (orange) is occluded both by the translocation domain belt and the long axis of the translocation domain (green).

The conserved residues of the catalytic domain were mapped onto the three-dimensional structure of BoNT/A and are shown in Figure 6. Conserved residues within 6 Å of the active site cavity are depicted with their full sidechains. In addition to the anticipated residues of the HEXXH motif (H222, E223, and H226), both E260 and E261 (where E261 can coordinate the zinc) and R362 and Y365 of loop D are located within this proximity. These seven residues are located in a plane at the base of the active site channel thus making it reasonable to assume that their conservation preserves the general zinc endopeptidase activity. The specificity is therefore likely to arise from the residues forming the channel, or cavity, above this base. This cavity is at least 1075 $Å^3$ in the presence of the translocation domain. Non-conserved catalytic domain residues within 4 Å of this cavity include 65-68 of loop B, 161-163, 193, 219, 238 of loop C, 255-258, and 368-9 of loop D. These residues were analyzed in smaller subsets, particularly BoNT/B and TeNT, for sequence conservation, but none was readily apparent. Residues 459, 530, 531, 533, and 536 of the translocation domain also border the active site cavity where P531 is conserved in all 10 sequences and the only strictly conserved residue of the belt. These residues, which occlude access to the active site in the holotoxin form, are presumably absent following translocation through the membrane. The possibility that the catalytic domain can adopt an altered conformation following its translocation awaits a structure determination of this domain in the absence of the holotoxin.

Figure 6: The BoNT/A catalytic domain colored to show residues with high sequence conservation in purple. Seven strictly conserved residues within 6 Å of the active site zinc (gray sphere) are depicted in orange and include H222, E223, H226, E260, E261, R362, and Y365.

While some level of sequence conservation was expected in the vicinity of the active site, the appearance of conservation in two surface loops, distant from the active site was surprising. Loop A (1-18) is comprised of conserved residues F7, N8, Y9, D11, and P12 and is close to conserved residues of β1, β2, β3, β6, as well as the N-terminus of α1 (Figure 7). The importance of this loop was identified in experiments measuring the toxicity of N-terminal and C-terminal catalytic domain deletions (Kurazono et al., 1992). While a deletion of residues 1-7 was tolerated, the deletion of residues 1-9 was not, implicating a direct biochemical and/or indirect structural role for one or both of residues N8 and Y9. While N8 is largely solvent accessible, Y9 forms part of a conserved charged pocket containing K33 (β2), E46 (β3), D80 (α1), and K83 (α1) (Figure 7). Since D80 and K83 are contained within one end of α1 and the other end of α1 has contact with α3, the HEXXH helix, it is possible that this conservation is merely to preserve the orientation of these helices, and therefore, the structural integrity of the catalytic machinery. This possibility seems unlikely however given the 30 Å distance of Y9 from the zinc and the lack of structural conservation in secondary structural elements more proximal to the HEXXH helix (for

example, β7 and β8). It seems more likely that the conserved presence of Tyr in the charged pocket is preserving a remote tertiary structure critical for catalysis. This is supported by experiments showing that a monoclonal antibody (mAB) mapping to residues 27-52, when administered intraneurally, neutralized the effects of BoNT/A (Cenci Di Bello *et al.*, 1994). The solvent accessible residues of this sequence and likely mAB binding site immediately preceed and follow the strands β2 and β3, respectively.

Figure 7: The BoNT/A catalytic domain is shown in purple with residues 1-88 depicted in yellow. Residues 450-550 of the translocation domain are colored green to show the relative position of the translocation domain 'belt'. The Asp-Pro residues are depicted in red while the conserved cluster of residues around Tyr9 is shown in purple. The backbone corresponding to the three Prolines P60-P62 is colored in green and shows the relative proximity to the active site zinc (gray sphere). Labels are underneath their corresponding secondary structural element.

The second conserved surface loop, the last part of loop B, is also located within this proximity. It is intriguing that both loops A and B contain an Asp-Pro pair of residues, although the conservation in loop B did not extend to BoNT/A (Figure 7). While acidic pH is not required for catalysis, cleavage in these sites may aid the efficiency of translocation. The conservation in the latter part of loop B is also of interest as the preceeding residues come into contact with the active site cavity. An unusual sequence of three consecutive prolines (P60, P61, and P62) is followed by residues 65-68 (previously noted for their proximity to the channel).

There is mounting evidence suggesting the existence of a remote recognition sequence necessary for catalysis. A ten residue sequence shared by all three of the SNARE proteins has been identified and termed the SNARE motif (Rossetto *et al.*, 1994). This motif is thought to be required for efficient catalysis and is present at varying distances upstream of the cleavage site. A potentially related study shows that the addition of peptides in trans upstream and downstream of the TeNT cleavage site is able to activate the toxin (Cornille *et al.*, 1997). The idea of exosite-dependent cleavage implies that a conformational change, in addition to simple removal of the translocation domain, may be required for catalysis. Perhaps the regions in BoNT/A that share high degrees of sequence conservation distal from the active site represent such general activation domains. Alternatively, these conserved surface loops could represent oligomeric interfaces or be involved in localization of the toxin within the cytosol.

8) Isolated Domains of Botulinum Neurotoxin Serotype A

Recombinant DNA work on serotype A domains to produce protein for crystallography experiments. High expression levels of the translocation domain of BT serotype A have been achieved and purified. The E. coli expressed protein is folded and appears to be composed of both a-helices and beta-sheets based on CD (circular dichroism) experiments conducted in my laboratory. Since the present crystals of botulinum neurotoxin diffract to 3.0 Å resolution, we will be able to determine the 3-dimensional structure of the neurotoxin and observe secondary structure elements (alpha helices, beta strands, and beta sheets). It will be difficult to observe detailed side chain interactions of the smaller side chains (the larger side chains should easily be observable). In order to obtain a more detailed picture of the protein structure, recombinant DNA work on isolated domains (binding, translocation, catalytic) have been initiated with the goal of crystallizing the domains and determining the 3-D structure by x-ray crystallography. The approach of "divide & conquer" has been used in numerous examples to determine the structure of regions of protein molecules (i.e. SH2 & SH3 domains of tyrosine kinases).

All work on this aspect of the project has been conducted by personnel supported by the Department of Chemistry at UC-Berkeley. Proper authorization to conduct recombinant DNA work on fragments of botulinum neurotoxin were obtained from the Biosafety Officer, Office of Environment, Health and Safety (see Annual Report from previous year).

This work is in line with goals C5 and C7 of the contracted work, to crystallize the isolated domains of botulinum neurotoxin. Large quantities of purified protein is required and a recombinant approach is being taken. The alternative approach is by purifying the light and heavy chains from one another. This approach would require twice or greater the amount of holo-neurotoxin plus purification. Secondly, it is impossible to separate the binding domain from the translocation domain in this fashion.

9) Proteolytic Assay of Catalytic Domain

Because of the need for a means to evaluate the quality of protein used in crystal growth experiments in addition to the fact that we need to learn as much as possible about the behavior of the protein during crystal growth experiments and to help us understand the function of the neurotoxin once the structure is determined, we have initiated simple kinetic assays to evaluate the neurotoxins proteolytic activity. The substrate being used is a 17 amino acid peptide of the synaptic vesicle protein SNAP-25. After several months of attempting to optimize proteolytic activity, it was determined that the zinc concentration of the toxin must be 100 micromolar. If the concentration of any zinc salt is above the micromolar range, the neurotoxin activity is inhibited. If the concentration of any zinc salt is below the micromolar range, the neurotoxin activity is very low (less than 10% of full activity). With this information, we now routinely use 100 •M zinc acetate in all crystallization conditions to help stabilize the neurotoxin. It cannot be overemphasized that micromolar concentrations of zinc are absolutely necessary for full activity. In similar zinc protease structures, it is observed that correct concentrations of zinc atoms can possess either structural/catalytic or both roles in aiding the proteolytic activity. In the case of botulinum neurotoxin, we believe zinc to play both roles since we have observed that zinc salts stabilize the crystal growth under the optimal conditions (micromolar).

Analogous to other crystal structure determinations, it has frequently been observed that inhibitors for enzymes aid in the stabilization of protein molecules by locking the enzyme into a single stable conformation. By using the above kinetic assay, we have been able to screen potential inhibitors of the neurotoxin to stabilize a single conformation. The inhibitors being investigated including peptide-like analogs synthesized by Professor Paul Bartlett in the Dept. of Chemistry of UC-Berkeley. Based on the protease recognition site, we are investigating the inhibitors Z-Ala-Gly-P-Phe-, Z-Phe-P-Leu-Ala-, Z-Gly-Ala-P-, and CbZ-Gly-P-Leu-Gly-. The "-P-" moiety is a phosphate backbone in place of the amide backbone that strengthens the peptide bond and does not allow cleavage by the protease. The inhibitor does however have the recognition elements that the neurotoxin binds.

Relationship to Goals of Research

All work to date focuses on the single goal of determining the 3-D structure. Although numerous side projects have emerged (kinetic assay, binding assays, recombinant DNA work), all of these projects are to aid in the crystal structure determination goal. Furthermore, the information obtained during these studies will be mandatory in understanding the function of the neurotoxin once the structure is known. All of the side projects listed above have been completed by personnel not supported by this contract. The individuals supported by the contract work only on the crystal structure determination.

REFERENCES

Aureli, P., Fenicia, L., Pasolini, B., Gianfranceschi, M., McCroskey, L. M. & Hatheway, C. L. (1986). Two cases of type E infant botulism caused by neurotoxigenic *Clostridium butyricum* in Italy. *J Infect Dis* **154**, 207-211.

Binz, T., Blasi, J., Yamasaki, S., Baumeister, A., Link, E., Sudhof, T. C., Jahn, R. & Niemann, H. (1994). Proteolysis of SNAP-25 by types E and A botulinal neurotoxins. *J Biol Chem* **269**(3), 1617-20.

Black, J. D. & Dolly, J. O. (1986a). Interaction of 125I-labeled botulinum neurotoxins with nerve terminals. I. Ultrastructural autoradiographic localization and quantitation of distinct membrane acceptors for types A and B on motor nerves. *J Cell Biol* **103**(2), 521-34.

Black, J. D. & Dolly, J. O. (1986b). Interaction of 125I-labeled botulinum neurotoxins with nerve terminals. II. Autoradiographic evidence for its uptake into motor nerves by acceptor-mediated endocytosis. *J Cell Biol* **103**(2), 535-44.

Blasi, J., Chapman, E. R., Link, E., Binz, T., Yamasaki, S., De Camilli, P., Sudhof, T. C., Niemann, H. & Jahn, R. (1993a). Botulinum neurotoxin A selectively cleaves the synaptic protein SNAP-25. *Nature* **365**(6442), 160-3.

Blasi, J., Chapman, E. R., Yamasaki, S., Binz, T., Niemann, H. & Jahn, R. (1993b). Botulinum neurotoxin C1 blocks neurotransmitter release by means of cleaving HPC-1/syntaxin. *Embo J* **12**(12), 4821-8.

Blaustein, R. O., Germann, W. J., Finkelstein, A. & DasGupta, B. R. (1987). The N-terminal half of the heavy chain of botulinum type A neurotoxin forms channels in planar phospholipid bilayers. *FEBS Lett* **226**(1), 115-20.

Cenci Di Bello, I., Poulain, B., Shone, C. C., Tauc, L. & Dolly, J. O. (1994). Antagonism of the intracellular action of botulinum neurotoxin type A with monoclonal antibodies that map to light-chain epitopes. *Eur J Biochem* **219**(1-2), 161-9.

Chen, F., Kuziemko, G. M., Amersdorfer, P., Wong, C., Marks, J. D. & Stevens, R. C. (1997). Antibody mapping to domains of botulinum neurotoxin serotype A in the complexed and uncomplexed forms. *Infect Immun* **65**(5), 1626-30.

Colman, P. M., Jansonius, J. N. & Matthews, B. W. (1972). The structure of thermolysin: an electron density map at 2-3 A resolution. *J Mol Biol* **70**(3), 701-24.

Cornille, F., Martin, L., Lenoir, C., Cussac, D., Roques, B. P. & Fournie-Zaluski, M. C. (1997). Cooperative exosite-dependent cleavage of synaptobrevin by tetanus toxin light chain. *J Biol Chem* **272**(6), 3459-64.

DasGupta, B. R. & Evenson, M. L. (1992). Botulinum neurotoxin type A: The isolated light chain breaks down. *FASEB Journal* **6**(1), A227.

DasGupta, B. R. & Tepp, W. (1991). Botulinum neurotoxin type A breaks down at pH 5.0 at the two Asp-Pro bonds of the heavy chain. *Society for Neuroscience Abstracts* **17**, 1526.

Dolly, J. O., Black, J., Williams, R. S. & Melling, J. (1984). Acceptors for botulinum neurotoxin reside on motor nerve terminals and mediate its internalization. *Nature* **307**(5950), 457-60.

Donovan, J. J. & Middlebrook, J. L. (1986). Ion-conducting channels produced by botulinum toxin in planar lipid membranes. *Biochemistry* **25**(10), 2872-6.

Foran, P., Lawrence, G. W., Shone, C. C., Foster, K. A. & Dolly, J. O. (1996). Botulinum neurotoxin C1 cleaves both syntaxin and SNAP-25 in intact and permeabilized chromaffin cells: correlation with its blockade of catecholamine release. *Biochemistry* **35**(8), 2630-6.

Hall, J. D., McCroskey, L. M., Pincomb, B. J. & Hatheway, C. L. (1985). Isolation of an organism resembling Clostridium barati which produces type F botulinal toxin from an infant with botulism. *J Clin Microbiol* **21**(4), 654-5.

Henikoff, J. G. & Henikoff, S. (1996). Blocks database and its applications. *Methods Enzymol* **266**, 88-105.

Hoch, D. H., Romero-Mira, M., Ehrlich, B. E., Finkelstein, A., DasGupta, B. R. & Simpson, L. L. (1985). Channels formed by botulinum, tetanus, and diphtheria toxins in planar lipid bilayers: relevance to translocation of proteins across membranes. *Proc Natl Acad Sci U S A* **82**(6), 1692-6.

Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**(12), 2577-637.

Kamata, Y., Yoshimoto, M. & Kozaki, S. (1997). Interaction between botulinum neurotoxin type A and ganglioside: ganglioside inactivates the neurotoxin and quenches its tryptophan fluorescence. *Toxicon* **35**(8), 1337-40.

Kraulis, P. J. (1991). MOLSCRIPT- a program to produce both detailed and schematic plots of protein structures. *J. Applied Crystallogr.* **24**, 946-950.

Kubota, T., Watanabe, T., Yokosawa, N., Tsuzuki, K., Indoh, T., Moriishi, K., Sanda, K., Maki, Y.,

Inoue, K. & Fujii, N. (1997). Epitope regions in the heavy chain of Clostridium botulinum type E neurotoxin recognized by monoclonal antibodies. *Appl Environ Microbiol* **63**(4), 1214-8.

Kurazono, H., Mochida, S., Binz, T., Eisel, U., Quanz, M., Grebenstein, O., Wernars, K., Poulain, B., Tauc, L. & Niemann, H. (1992). Minimal essential domains specifying toxicity of the light chains of tetanus toxin and botulinum neurotoxin type A. *J Biol Chem* **267**(21), 14721-9.

Lacy, D. B. & Stevens, R. C. (1998). Unraveling the structures and modes of action of bacterial toxins. *Curr Opin Struct Biol* **8**(6), 778-84.

Lacy, D. B., Tepp, W., Cohen, A. C., DasGupta, B. R. & Stevens, R. C. (1998). Crystal structure of botulinum neurotoxin type A and implications for toxicity. *Nat Struct Biol* **5**(10), 898-902.

Lebeda, F. J. & Olson, M. A. (1994). Secondary structural predictions for the clostridial neurotoxins. *Proteins* **20**(4), 293-300.

Lebeda, F. J. & Olson, M. A. (1995). Structural predictions of the channel-forming region of botulinum neurotoxin heavy chain. *Toxicon* **33**(4), 559-67.

Merritt, E. A. & Bacon, D. J. (1997). Raster3D: Photorealistic molecular graphics. *Meth. Enz.* **277**, 505-524.

Montal, M. S., Blewitt, R., Tomich, J. M. & Montal, M. (1992). Identification of an ion channel-forming motif in the primary structure of tetanus and botulinum neurotoxins. *FEBS Lett* **313**(1), 12-8.

Montecucco, C. (1986). How do tetanus and botulinum toxins bind to neuronal membranes? *Trends Biochem Sci* **11**, 315-317.

Montecucco, C. & Schiavo, G. (1995). Structure and function of tetanus and botulinum neurotoxins. *Q Rev Biophys* **28**(4), 423-72.

Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Prot. Struct. Funct. Genet.* **11**, 281-296.

Oblatt-Montal, M., Yamazaki, M., Nelson, R. & Montal, M. (1995). Formation of ion channels in lipid bilayers by a peptide with the predicted transmembrane sequence of botulinum neurotoxin A. *Protein Sci* **4**(8), 1490-7.

Risler, J. L., Delorme, M. O., Delacroix, H. & Henaut, A. (1988). Amino acid substitutions in structurally related proteins. A pattern recognition approach. Determination of a new and efficient scoring matrix. *J Mol Biol* **204**(4), 1019-29.

Rossetto, O., Schiavo, G., Montecucco, C., Poulain, B., Deloye, F., Lozzi, L. & Shone, C. C. (1994). SNARE motif and neurotoxins. *Nature* **372**(6505), 415-6.

Schiavo, G., Benfenati, F., Poulain, B., Rossetto, O., Polverino de Laureto, P., DasGupta, B. R. & Montecucco, C. (1992). Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic cleavage of synaptobrevin. *Nature* **359**(6398), 832-5.

Schiavo, G., Malizio, C., Trimble, W. S., Polverino de Laureto, P., Milan, G., Sugiyama, H., Johnson, E. A. & Montecucco, C. (1994). Botulinum G neurotoxin cleaves VAMP/synaptobrevin at a single Ala-Ala peptide bond. *J Biol Chem* **269**(32), 20213-6.

Schiavo, G., Rossetto, O., Catsicas, S., Polverino de Laureto, P., DasGupta, B. R., Benfenati, F. & Montecucco, C. (1993a). Identification of the nerve terminal targets of botulinum neurotoxin serotypes A, D, and E. *J Biol Chem* **268**(32), 23784-7.

Schiavo, G., Santucci, A., Dasgupta, B. R., Mehta, P. P., Jontes, J., Benfenati, F., Wilson, M. C. & Montecucco, C. (1993b). Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct COOH-terminal peptide bonds. *FEBS Lett* **335**(1), 99-103.

Schiavo, G., Shone, C. C., Rossetto, O., Alexander, F. C. & Montecucco, C. (1993c). Botulinum neurotoxin serotype F is a zinc endopeptidase specific for VAMP/synaptobrevin. *J Biol Chem* **268**(16), 11516-9.

Schmid, M. F., Robinson, J. P. & DasGupta, B. R. (1993). Direct visualization of botulinum neurotoxin-induced channels in phospholipid vesicles. *Nature* **364**(6440), 827-30.

Shapiro, R. E., Specht, C. D., Collins, B. E., Woods, A. S., Cotter, R. J. & Schnaar, R. L. (1997). Identification of a ganglioside recognition domain of tetanus toxin using a novel ganglioside photoaffinity ligand. *J Biol Chem* **272**(48), 30380-6.

Sheridan, R. E. (1998). Gating and permeability of ion channels produced by botulinum toxin types A and E in PC12 cell membranes. *Toxicon* **36**(5), 703-17.

Shone, C. C., Hambleton, P. & Melling, J. (1985). Inactivation of Clostridium botulinum type A neurotoxin by trypsin and purification of two tryptic fragments. Proteolytic action near the COOH-terminus of the heavy subunit destroys toxin-binding activity. *Eur J Biochem* **151**(1), 75-82.

Simpson, L. L. (1980). Kinetic studies on the interaction between botulinum toxin type A and the cholinergic neuromuscular junction. *J Pharmacol Exp Ther* **212**(1), 16-21.

Simpson, L. L. (1989). *Botulinum Neurotoxin and Tetanus Toxin*, Academic Press, San Diego.

Stevens et al,*J.*(1991) Crystallization of Botulinum Neurotoxin Serotype A *Mol. Biol.* **222**, 877.

Sutton, R. B., Fasshauer, D., Jahn, R. & Brunger, A. T. (1998). Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 A resolution. *Nature* **395**(6700), 347-53.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**(22), 4673-80.

Umland, T. C., Wingert, L. M., Swaminathan, S., Furey, W. F., Schmidt, J. J. & Sax, M. (1997). Structure of the receptor binding fragment HC of tetanus neurotoxin. *Nat Struct Biol* **4**(10), 788-92.

Yamasaki, S., Binz, T., Hayashi, T., Szabo, E., Yamasaki, N., Eklund, M., Jahn, R. & Niemann, H. (1994a). Botulinum neurotoxin type G proteolyses the Ala81-Ala82 bond of rat synaptobrevin 2. *Biochem Biophys Res Commun* **200**(2), 829-35.

Yamasaki, S., Hu, Y., Binz, T., Kalkuhl, A., Kurazono, H., Tamura, T., Jahn, R., Kandel, E. & Niemann, H. (1994b). Synaptobrevin/vesicle-associated membrane protein (VAMP) of Aplysia californica: structure and proteolysis by tetanus toxin and botulinal neurotoxins type D and F. *Proc Natl Acad Sci U S A* **91**(11), 4688-92.

CONCLUSION

We have accomplished the primary objective outlined in the research contract - the three-dimensional structure determination of botulinum neurotoxin serotype A. We have cloned and expressed the isolated domains, and crystallized the translocation domain. We have crystallized two other serotypes of botulinum neurotoxin (serotype B and E). This work has opened the door to a great deal of future work on understanding how the molecule functions. Furthermore, the results open the door for the design of inhibitors for the toxins activity. Already, we can speculate how the translocation domain interacts with the membrane based on the long helices analogous to the helices in influenza virus hemaglutinin.

Future work on this project includes an increase in diffraction resolution for a higher resolution structure, experiments to deduce the mechanism of action for each of the domains (binding, translocation, catalytic) of botulinum neurotoxin, and structure determination of the other serotypes and isolated domains.

BIBLIOGRAPHY

Publications

1. F. Chen, G. Kuziemko, P. Amersdorfer, C. Wong, J.D. Marks, R.C. Stevens "Antibody Mapping to Domains of Botulinum Neurotoxin serotype A in the Complexed and Uncomplexed Forms" *Infection & Immunity* **65**, 1626 (1997).

2. R. Burkhard, F. Chen, G. Kuziemko, R.C. Stevens "Electron Density Projection Maps of the 900 kD Botulinum Neurotoxin Complex by Electron Crystallography" *J. Structural Biology* **120**, 78 (1997).

3. D.B. Lacy and R.C. Stevens "Recombinant Expression and Purification of the Botulinum Neurotoxin A Translocation Domain", *Protein Expression and Purification* **11**, 195 (1997).

4. F. Chen, G. M. Kuziemko, R.C. Stevens "Biophysical Characterization of the Stability of the 150-Kilodalton Botulinum Toxin, the Non-toxic Component, and the 900-Kilodalton Botulinum Toxin Complex Species" *Infection & Immunity* **66**, 2420 (1998).

5. B. Lacy, W. Tepp, B. DasGupta, R.C. Stevens "Three-dimensional Crystal Structure of Botulinum Neurotoxin and Implications in Toxicity" *Nature Structural Biology* **5**, 898 (1998).

6. B. Lacy and R.C. Stevens, "Unraveling Structures and Modes of Action of Bacterial Toxins"*Current Opinions Structural Biology* **8**, 778 (1998).

7. B. Lacy and R.C. Stevens, "Sequence homology and structural analysis of the clostridial neurotoxins", in press *J. Mol. Biol.* (1999).

8. F. Chen, G. Kuziemko, R.C. Stevens, "Structure of the 900 kDa Botulinum Neurotoxin Complex and Comparison to the 3D holotoxin X-ray Structure", submitted *Structure* (1999).

Student Theses

Kuziemko, G.M. "A Biophysical Characterization of Botulinum Neurotoxin and Cholera Toxin", Ph.D. Thesis, Department of Chemistry, University of California, Berkeley, 1997.

Chen, F. "Botulinum Neurotoxin Complex: A Structural and Biophysical Characterization", Ph.D. Thesis, Biophysics Group, University of California, Berkeley, 1998.

Lacy, D.B. "A Structural Study of Botulinum Neurotoxin" Ph.D. Thesis, Department of Chemistry, University of California, Berkeley, 1999.

Meeting Abstracts

"Structural Features of Botulinum Neurotoxin Serotype A toxicity" 2nd International Meeting on the Molecular Genetics and Pathogenesis of the Clostridia, Onzain, France June 22-25, 1997.

1997 IBRCC Meeting, Washington, D.C.

1998 IBRCC Meeting, Pennsylvania

Personnel Conducting Research on the Project at Univ. California-Berkeley

Professor Raymond C. Stevens
Dr. Cara Marks
Dr. Hanne Merritt
Heather Manning
Borden Lacy
Alona Cohen

Personnel Conducting Research on the project at U. Wisconsin-Madison (sub-contract)

Bibhuti DasGupta
Bill Tepp
Mary Evenson

APPENDIX (Peer Reviewed Published Manuscripts)

1. F. Chen, G. Kuziemko, P. Amersdorfer, C. Wong, J.D. Marks, R.C. Stevens "Antibody Mapping to Domains of Botulinum Neurotoxin serotype A in the Complexed and Uncomplexed Forms" *Infection & Immunity* **65**, 1626 (1997).

2. R. Burkhard, F. Chen, G. Kuziemko, R.C. Stevens "Electron Density Projection Maps of the 900 kD Botulinum Neurotoxin Complex by Electron Crystallography" *J. Structural Biology* **120**, 78 (1997).

3. D.B. Lacy and R.C. Stevens "Recombinant Expression and Purification of the Botulinum Neurotoxin A Translocation Domain", *Protein Expression and Purification* **11**, 195 (1997).

4. F. Chen, G. M. Kuziemko, R.C. Stevens "Biophysical Characterization of the Stability of the 150-Kilodalton Botulinum Toxin, the Non-toxic Component, and the 900-Kilodalton Botulinum Toxin Complex Species" *Infection & Immunity* **66**, 2420 (1998).

5. B. Lacy, W. Tepp, B. DasGupta, R.C. Stevens "Three-dimensional Crystal Structure of Botulinum Neurotoxin and Implications in Toxicity" *Nature Structural Biology* **5**, 898 (1998).

6. B. Lacy and R.C. Stevens, "Unraveling Structures and Modes of Action of Bacterial Toxins"*Current Opinions Structural Biology* **8**, 778 (1998).

7. B. Lacy and R.C. Stevens, "Sequence homology and structural analysis of the clostridial neurotoxins", in press *J. Mol. Biol.* (1999).

8. F. Chen, G. Kuziemko, R.C. Stevens, "Structure of the 900 kDa Botulinum Neurotoxin Complex and Comparison to the 3D holotoxin X-ray Structure", submitted *Structure* (1999).

INFECTION AND IMMUNITY, May 1997, p. 1626–1630
0019-9567/97/$04.00+0

Vol. 65, No. 5

Antibody Mapping to Domains of Botulinum Neurotoxin Serotype A in the Complexed and Uncomplexed Forms

FLORA CHEN,[1] GEOFFREY M. KUZIEMKO,[2] PETER AMERSDORFER,[3] CINDY WONG,[3] JAMES D. MARKS,[3]* AND RAYMOND C. STEVENS[1,2]*

Graduate Group in Biophysics[1] and Department of Chemistry,[2] University of California, Berkeley, California 94720, and Departments of Anesthesia and Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94110[3]

Received 8 November 1996/Returned for modification 13 January 1997/Accepted 7 February 1997

The domain organization of the botulinum neurotoxin serotype A was studied by using antibody mapping of 44 monoclonal single-chain variable fragments. The analysis was carried out on (i) the individual domains of botulinum neurotoxin holotoxin (binding, translocation, and catalytic), (ii) botulinum neurotoxin holotoxin, (iii) the botulinum neurotoxin holotoxin in complex with the nontoxic portion, and (iv) botulinum neurotoxin holotoxin and nontoxic portion of the complex recombined in vitro. All 44 antibodies mapped to individual domains of botulinum neurotoxin. Forty of the 44 single-chain variable fragments bound the botulinum neurotoxin holotoxin relative to the isolated domains, suggesting that 4 epitopes are covered when the individual domains are in the holotoxin form. Only 20 of the antibodies showed a positive reaction to the toxin while in complex with the nontoxic portion. All of the covered epitopes were mapped to the binding domain of botulinum neurotoxin, which suggested that the binding domain is in direct contact with the nontoxic portion in the complex. Based on the antibody mapping to the different domains of the botulinum neurotoxin holotoxin and the entire complex, a model of the botulinum neurotoxin complex is proposed.

The anaerobic bacterium *Clostridium botulinum* produces seven serotypes of neurotoxin, classified A through G (26). The neurotoxin, serotype A, can be purified as a 900-kDa complex consisting of a 150-kDa toxic component (botulinum neurotoxin [BT]), and a 750-kDa nontoxic component (hemagglutinin [HA]) (3). The BT inhibits cholinergic vesicle docking at the neuromuscular junction, resulting in flaccid paralysis (23), and most commonly intoxicates by oral ingestion. The three 50-kDa functional domains of BT—binding (13), translocation (2), and catalytic (1a)—allow the toxin to bind to a cell surface receptor, pass across the membrane (23), and cleave a protein involved in vesicle docking, respectively (9). Sugii and coworkers (24) have shown that the HA-BT complex has a higher oral toxicity in rats than the BT alone. The 750-kDa HA has been shown to have an agglutination ability (15) and is thought to protect the toxin from the extreme pH and proteases in the gut (25).

Little structural information is known about BT, HA-BT, or the interaction between BT and HA. The literature contains examples of antibodies used to detect various serotypes of BT (7, 12, 14, 27), but no mapping of the BT domains in the HA-BT complex has been carried out. Sugiyama and coworkers (27) showed that polyclonal antibodies to type A HA-BT complex recognized epitopes predominantly from the HA and not the BT. Monoclonal antibodies developed by Kozaki and coworkers recognized the light chain of the BT that causes infant botulism and the light chain of BT serotype A (14). However, the antibodies to the infant botulism BT heavy chain did not recognize the BT serotype A heavy chain. Studies using polyclonal antibodies to the toxin detected BT at very low concentrations but did not provide specific information about the relationship between the BT and HA (7, 12). We probed the structure of the HA-BT complex by using a panel of 44 unique monoclonal single-chain variable fragments (scFv) derived from combinatorial phage antibody libraries (1, 28).

We used enzyme-linked immunosorbent assays (ELISAs) to identify 44 scFv that bind to different domains of botulinum neurotoxin serotype A. ELISAs were performed on purified BT, the purified BT domains, HA, HA-BT complex, and recombined HA-BT in vitro. Based on our results, we propose a model to illustrate the interaction between BT and HA. The model could act as a guide for the design of neutralizing antibodies and may explain how the HA protects the BT from proteolytic and pH attack.

MATERIALS AND METHODS

Purification of HA-BT complex, BT, and BT domains. The HA-BT complex (Hal strain) was obtained as an ammonium sulfate precipitate from purified bacterial supernatant at a concentration of 3.3 mg/ml in 50 mM sodium citrate (pH 5.5) (5). Before use, the HA-BT complex was centrifuged at 26,890 × *g* (Dupont Sorvall RC-5B centrifuge) for 15 min and dialyzed against saline (0.68 M sodium citrate, 0.145 M NaCl [pH 7.4]) with three buffer exchanges within an hour. Concentration was determined by A_{278} measurements (1.66 arbitrary units/mg ml^{-1}), using a Shimadzu UV-160 spectrophotometer (11).

The HA was purified in two steps by using a modification of a published procedure (4). Forty-five milligrams of ammonium sulfate precipitate of HA (0.42 g/ml) was centrifuged at 26,890 × *g*. The pellet was dissolved in 20 ml of 70 mM Tris-HCl (pH 7.2) and dialyzed overnight. The dialyzed solution was centrifuged at 26,890 × *g* for 15 min and applied onto a DEAE-Sepharose column (1.5 by 24 cm; Pharmacia, Uppsala, Sweden) that was equilibrated with 70 mM Tris-HCl (pH 7.2). The column was washed with 100 ml of 70 mM Tris-HCl (pH 7.2) and the HA was eluted with 70 mM Tris-HCl–0.2 M NaCl (pH 7.2). The fractions containing HA were combined and run on an SP-Sepharose column (1.5 by 21 cm) (Pharmacia, Uppsala, Sweden) equilibrated with 70 mM Tris-HCl (pH 7.2). Since residual BT adheres to the column matrix at this pH, the HA was collected in the flowthrough. The concentration of protein was determined by A_{278} measurements (11).

The BT was purified as described previously (5) and stored as a 10-mg/ml solution in 10 mM HEPES (pH 7.0)–0.1 M KCl–2 mM sodium azide. The binding domain of BT type A, expressed in *Escherichia coli* and purified by immobilized metal affinity chromatography using a C-terminal His_6 tag, was

* Corresponding author. Mailing address for James D. Marks: Departments of Anesthesia and Pharmaceutical Chemistry, University of California, San Francisco, SFGH Room 3C-38, 1001 Potrero, San Francisco, CA 94110. Phone: (415) 206-3256. Fax: (415) 206-3253. E-mail: jim_marks@quickmail.ucsf.edu. Phone for R. C. Stevens: (510) 643-8285. Fax: (510) 643-9290. E-mail: Stevens@adrenaline.berkeley.edu.

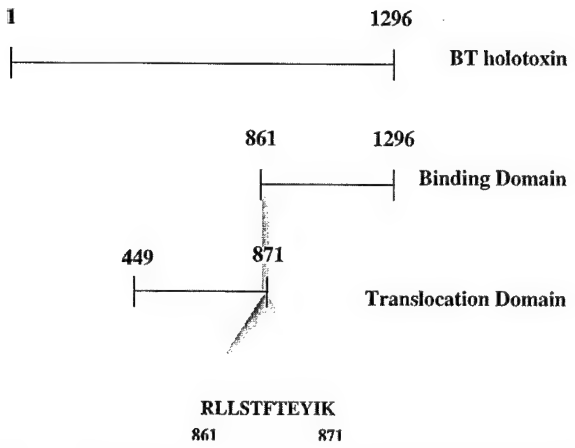


FIG. 1. Diagram illustrating the sequence overlap between the translocation and binding domain constructs. The top line represents the sequence of BT from the N terminus (residue 1) to the C terminus (residue 1296). The bottom two lines depict the C-terminal 11 residues from the translocation domain construct overlapping the N-terminal 11 residues of the binding domain (18, 19).

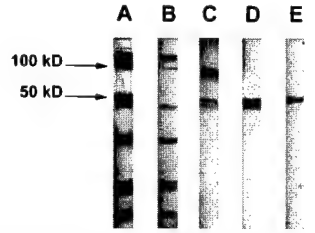


FIG. 2. SDS-PAGE analyses of column chromatography. Lanes: A, HA-BT complex; B, purified HA; C, purified BT; D, purified binding domain; E, purified translocation domain.

purchased from Ophidian Pharmaceuticals, Inc. (Madison, Wis.). The translocation domain (residues 449 to 872) of BT type A was expressed in *E. coli* and purified by immobilized metal affinity chromatography using a C-terminal His_6 tag (14a). The 11 C-terminal residues of the translocation domain overlap the 11 N-terminal residues of binding domain (18, 19) (Fig. 1). The polypeptide composition of each of the different batches of purified protein was analyzed on 12% polyacrylamide gels as described by Fling and Gregerson (8).

Antibodies. ScFv antibody fragments were selected from four different combinatorial phage antibody libraries (1, 28). Briefly, scFv phage antibody libraries were constructed from the immunoglobulin heavy (V_H)- and light (V_L)-chain variable regions of mice immunized with purified holotoxin (library 1); mice immunized with binding domain (library 2); humans immunized with pentavalent botulinum toxoid (Centers for Disease Control and Prevention) (library 3); nonimmunized human volunteers (library 4). Libraries 1 and 3 were constructed in the vector pCANTAB5E (Pharmacia), and library 2 was constructed in pHEN-1 (10). Specific scFv were isolated by selecting the libraries on either holotoxin or binding domain immobilized on polystyrene or in solution. The specificity of the isolated antibodies for the holotoxin or the binding domain was confirmed by ELISA on the relevant antigen and a panel of irrelevant antigens (1, 28). The number of unique scFv was determined by *Bst*N1 fingerprinting, followed by DNA sequencing of the V_H and V_L genes. Additional unique scFv were isolated from a 9.7×10^9-member nonimmune library in pHEN-1 (10) constructed from human V_H and V_L genes.

For structural mapping of the HA-BT complex, BT, and BT domains by ELISA, native scFv was expressed from the appropriate phagemid in *E. coli* HB2151 (10). The amber codon between the scFv gene and gene 3 permits expression of native scFv in a nonsuppressor *E. coli* strain (HB2151). scFv binding was detected by using the epitope tag at the C terminus of the scFv (E tag for scFv in pCANTAB5E and Myc tag for scFv in pHEN-1). Since both the scFv in pHEN-1 and the translocation domain have a C-terminal Myc tag, ELISA on the translocation domain was performed with scFv fused to phage, and detection was achieved by using HRP/anti-M13 conjugate (Pharmacia).

Expression of native scFv (6) in the phagemid vectors (pHEN-1 and pCANTAB5E) was performed in 96-well microtiter plates as described previously (16), with the following exception: after overnight growth and expression at 25°C, 50 μl of 0.5% Tween 20 was added to each well and the plates were incubated for 4 h at 37°C with shaking to induce bacterial lysis and increase the concentration of scFv in the bacterial supernatant. Supernatants containing native scFv were used for ELISA. To prepare phage for ELISA, single ampicillin-resistant colonies were transferred into microtiter plate wells containing 100 μl of 2×YT medium (16 g of Bacto Tryptone, 10 g of Bacto Yeast Extract, 5 g of NaCl, 1 liter of deionized H_2O) supplemented with 1 mM ampicillin (Sigma, St. Louis, Mo.) and 0.1% glucose. After 3 h of growth at 37°C to an A_{600} of approximately 0.5 arbitrary unit, VCSM13 helper phage (2.5×10^8 phage particles) was added, and the cells were incubated for 1 h at 37°C. Subsequently, kanamycin was added to a final concentration of 25 μg/ml, and the bacteria were grown overnight at 37°C. Supernatants containing phage were used for ELISA.

ELISA. Microtiter plates (Falcon 3912) were incubated with 50 μl of antigen (10 μg/ml except for the binding domain, in which case 5 μg/ml was used) in phosphate-buffered saline (PBS; 25 mM NaH_2PO_4, 125 mM NaCl [pH 7.0]) at 4°C overnight. After being washed once with PBS, wells were incubated with 50 μl of bacterial supernatant containing either native scFv or scFv fused to phage. Myc-tagged scFv were detected with mouse monoclonal antibody 9E10 (1 μg/ml; Santa Cruz Biotechnology) (17), and E-tagged scFv were detected by using anti-E tag antibody (1 mg/ml; Pharmacia), followed by peroxidase-conjugated anti-mouse Fc antibody (Sigma), with 2,2′-azinobis(3-ethylbenzthiazoline-sulfonic acid) as the substrate as described previously (21). Binding of scFv phage to antigens was detected with peroxidase-conjugated anti-M13 antibody (Pharmacia) as described elsewhere (20).

Recombination. The purified BT and HA were incubated at a 1:1 molar ratio for at least 24 h in 70 mM Tris-HCl (pH 7.2) at a protein concentration of 0.1 mg/ml. The mixture was stored at 4°C until diluted in PBS for coating ELISA plates.

RESULTS

Purification of antigens. After column chromatography, the purities of HA-BT complex, BT, and HA were ascertained by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) to be relatively free of contaminants (Fig. 2). The BT appeared as two bands, separated due to the reducing environment of the gel loading buffer—a 50-kDa light chain consisting of catalytic domain and a 100-kDa heavy chain consisting of translocation and binding domains. The HA-BT complex runs as nine bands that make up the BT and large HA. Purified HA appears identical to the HA-BT complex minus the 50-kDa polypeptide from the light chain of BT. A small amount of 100-kDa band from the heavy chain of BT was present in the HA. This residual BT was not enough to produce a positive signal for HA on ELISA. Binding and translocation domains, expressed in *E. coli* were also tested for purity by SDS-PAGE.

Antibody isolation and initial characterization. scFv were isolated by selection of phage libraries on immobilized purified BT or purified recombinant binding domain. The antibodies recognize only native protein and not denatured protein. Thus, the antibodies bound according to structure and not sequence libraries (1, 28). After three rounds of selection, initial scFv characterization by ELISA on BT and DNA sequencing of the V_H and V_L genes yielded 44 unique scFv (Table 1).

Structural mapping of BT, HA-BT complex, and BT domains. The forty-four scFv recognized the individual domains of BT, and none bound to HA alone (Table 1). Of the 44 scFv, 24 mapped to the binding domain and 3 mapped to the translocation domain (Table 1). In addition, two antibodies mapped to both the binding and translocation domains. These two scFv presumably bound to the 11-amino-acid overlap at the C terminus of the translocation domain and the N terminus of the binding domain (Fig. 1). Alternatively, cross-reactivity could result from binding to a four-amino-acid sequence (-KYVD-) that is homologous for residues 855 to 858 of the translocation domain and 1121 to 1124 of the binding domain. Thus, 26 scFv recognized the binding domain and 5 bound the translocation domain construct. The remaining 15 scFv presumably bound to the catalytic domain, though this was not tested directly due to lack of purified catalytic domain. Some of these 15 may recognize epitopes that are shared between domains.

TABLE 1. ELISA absorbances of scFv to type A neurotoxin[a]

Antibody	OD_{405} (avg ± SD)				
	BT	HA-BT	Binding domain	Translocation domain	RR
3d12	2.303 ± 0.682	0.659 ± 0.072	2.260 ± 0.489		1.181 ± 0.303
3a6	2.053 ± 0.768	0.660 ± 0.083	1.775 ± 0.263		1.160 ± 0.322
3d4	0.964 ± 0.192	0.913 ± 0.192			1.091 ± 0.108
4a4	0.950 ± 0.135	0.905 ± 0.045			1.195 ± 0.071
3a2	1.091 ± 0.192	0.936 ± 0.138			1.204 ± 0.234
3e3	1.106 ± 0.170	0.987 ± 0.158			1.229 ± 0.255
3e8	1.058 ± 0.358	0.941 ± 0.097			1.145 ± 0.245
3a11	1.036 ± 0.298	0.934 ± 0.108			1.248 ± 0.157
3e7	1.022 ± 0.262	0.945 ± 0.156			1.218 ± 0.153
3h3	1.143 ± 0.396	0.838 ± 0.380			1.037 ± 0.316
3a1	1.019 ± 0.300	0.943 ± 0.148			1.206 ± 0.167
w3	2.083 ± 0.523	1.544 ± 0.192			1.925 ± 0.267
w42	1.796 ± 0.409	1.400 ± 0.223			1.721 ± 0.270
g23	1.875 ± 0.403	1.460 ± 0.194			1.795 ± 0.183
g3	1.287 ± 0.329	0.890 ± 0.298			1.290 ± 0.174
g11	1.200 ± 0.326	0.956 ± 0.192			1.285 ± 0.100
w7	1.552 ± 0.269	1.199 ± 0.273		1.182 ± 0.121	1.606 ± 0.393
w9	1.492 ± 0.376	1.239 ± 0.247		1.201 ± 0.152	1.510 ± 0.194
g7	1.072 ± 0.481	1.106 ± 0.327		0.590 ± 0.105	1.422 ± 0.261
w20	1.557 ± 0.530		1.713 ± 0.349		1.089 ± 0.053
w36	1.302 ± 0.418		1.362 ± 0.300		0.680 ± 0.026
w43	1.421 ± 0.253	1.002 ± 0.109			1.283 ± 0.334
g53	1.193 ± 0.321		1.700 ± 0.242		0.685 ± 0.095
g57	1.647 ± 0.409		1.423 ± 0.521		0.761 ± 0.117
c9	1.560 ± 0.460		1.690 ± 0.229		0.889 ± 0.213
c15	1.936 ± 0.535		1.988 ± 0.458		1.079 ± 0.358
s25	1.395 ± 0.363		1.369 ± 0.435		
3f6	1.332 ± 0.338		1.553 ± 0.272		0.811 ± 0.142
2a2	0.816 ± 0.297		1.172 ± 0.262		0.838 ± 0.033
2b10	0.858 ± 0.332		1.204 ± 0.408		0.661 ± 0.143
2b1	0.836 ± 0.121		1.247 ± 0.278	1.227 ± 0.505	0.566 ± 0.010
3e6	0.909 ± 0.109		1.190 ± 0.135		0.605 ± 0.165
2e6	0.838 ± 0.181		1.046 ± 0.171		
3d1	0.908 ± 0.192		1.009 ± 0.385		0.752 ± 0.072
2b6	0.652 ± 0.298		0.711 ± 0.213		
2h6	1.010 ± 0.326		1.261 ± 0.346		0.928 ± 0.004
2a8	0.994 ± 0.275		1.069 ± 0.398		0.908 ± 0.226
id5	0.738 ± 0.289		1.273 ± 0.228		0.716 ± 0.085
ie8	0.556 ± 0.339		0.921 ± 0.513	1.000 ± 0.405	0.654 ± 0.057
ig7	0.834 ± 0.174		1.113 ± 0.313		0.688 ± 0.077
3c3			0.621 ± 0.026		
2c3			0.740 ± 0.065		
3e2			0.427 ± 0.089		
3c5			0.743 ± 0.204		

[a] Values obtained with antigen coated at 10 μg/ml except for the binding domain, which was coated at 5 μg/ml. OD_{405}, optical density at 405 nm; RR, recombination of purified BT and HA. The antibodies were produced in the lab of James D. Marks. The values represent the averages of 13 plates coated with BT, 10 plates coated with HA-BT complex, 6 plates coated with binding domain, 6 plates coated with translocation domain, 5 plates coated with HA, and 4 plates coated with recombined BT and HA. The background was defined as the signal from plates containing antigen and primary and secondary antibodies. Values represent absorbances after background absorbance is subtracted. Blanks in columns represent absorbances below three times the background absorbance. These scFv were assumed not to bind the corresponding antigen. Since the ELISAs were performed using different batches of supernatant, the average absorbance of each plate was normalized to the average absorbance of the first plate for a particular antigen.

Forty scFv bound to the holotoxin (Table 1). Thus, four epitopes were covered when the individual domains came together to form the BT. Of these 40 scFv, 22 mapped to the binding domain and 3 mapped to the translocation domain. The remaining 15 scFv were deduced to recognize the catalytic domain (Table 1).

ELISA of scFv on the HA-BT complex permitted identification of BT epitopes which were inaccessible in the HA-BT complex. Twenty epitopes were covered when in the HA-BT complex. All of these covered epitopes were localized to the binding domain. Of the 22 scFv that mapped to the binding domain and BT, only 2 scFv bound to the HA-BT complex (Table 1). The three scFv that bound to translocation domain and BT also bound to the HA-BT complex (Table 1). The 15 scFv that mapped to catalytic domain in BT likewise bound to the HA-BT complex (Table 1).

DISCUSSION

Using antibodies to map the different domains of botulinum neurotoxin serotype A, we propose a model illustrating how the toxin may bind into the HA assembly. Forty-four scFv were produced to the BT and its domains. The antibodies specific to individual domains were used to map relative positions of

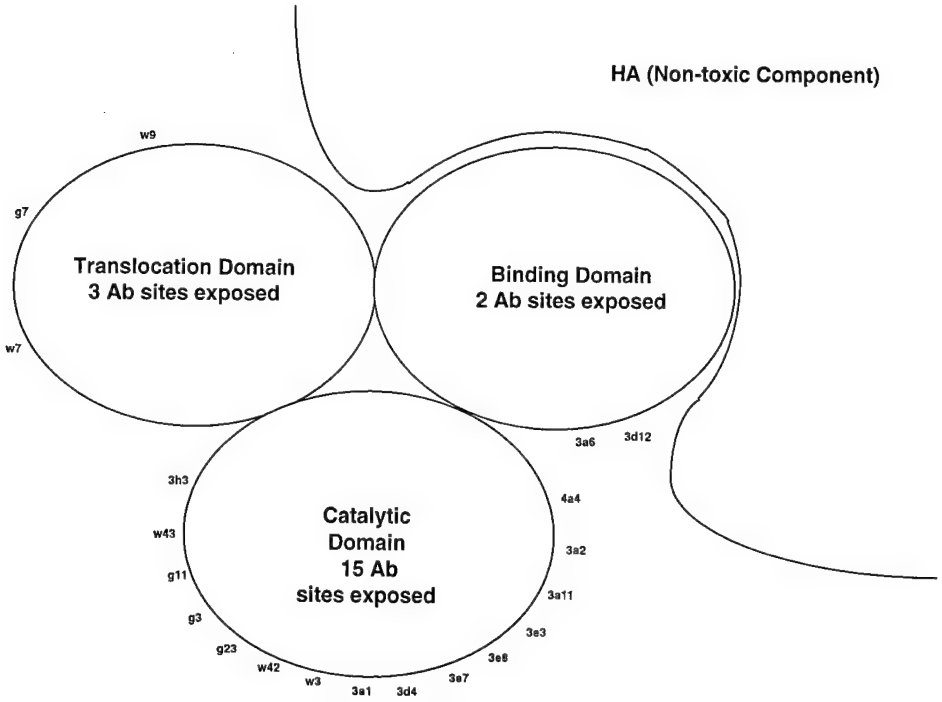


FIG. 3. Representation of a possible arrangement of BT with respect to HA. Only clones which bind the HA-BT complex are shown. The binding and translocation domains are covered partially by HA. Ab, antibody.

exposed or hidden epitopes in the BT and in the HA-BT complex. All of the antibodies specific to the translocation domain also recognized the HA-BT, indicating that the translocation domain is as exposed to the solvent in the complex (Fig. 3) as in the holotoxin (Table 1).

Not all of the scFv which bound the individual domains bound the BT holotoxin. Twenty-two of the 26 antibodies that recognized binding domain also recognized the BT. Thus, 4 antibody sites on the binding domain are covered by the catalytic or translocation domain in the BT (Table 1). However, not all of these scFv bound when the BT was complexed with the HA. Of the 22 exposed sites of binding domain on BT, only 2 bound the HA-BT complex. Therefore, the majority of antibodies to the binding domain recognize epitopes that must be covered by HA in the HA-BT complex (Fig. 3).

Fifteen antibodies were deduced to bind the catalytic domain. These antibodies were inferred by subtracting the antibodies that bound the binding and translocation domains from the antibodies that bound BT, since it is assumed that all of the antibodies must bind a domain of BT. All of these 15 antibodies bind an epitope that is accessible in the HA-BT complex. These 15 antibodies may bind conformational epitopes shared between domains. Such epitopes may not exist in the separate domains.

The model illustrated in Fig. 3 is subject to several caveats. Since the sequences of the epitopes are unknown, it is not possible to know the distribution of antibody binding sites. Hence the epitopes could be distributed evenly on the exposed surfaces or could be concentrated in certain regions of the protein. If the epitopes are clustered together, the area of BT covered by HA may be overestimated. If the epitopes are spaced regularly, the area of BT covered by HA may be underestimated. In either case, verification requires mapping of these antibodies to specific sequences of BT. Finally, since the model is drawn in two dimensions, it may not depict accurately the surface area of BT covered by HA. In addition, the arrangement of exposed and unexposed epitopes may be different from that diagrammed. The representation shows contiguous groups of exposed or unexposed epitopes that may be commingled.

The recombination experiments were performed to determine whether purified BT would interact with purified HA to reform the stable HA-BT complex at physiological pH. Since HA was undetectable by any of the antibody clones, the clones that bound to recombined HA-BT must bind exposed regions of BT. If no reconstitution of the complex occurred, recombined HA-BT should show the same number of clones as purified BT. If the HA-BT complex was formed, then recombined HA-BT should exhibit a positive reaction with the same clones as the purified HA-BT complex. The results from Table 1 indicate that an incomplete recombination took place. Whereas 40 clones recognized BT and 20 clones recognized the HA-BT complex, 37 clones bound recombined HA-BT. The three clones found in recombined HA-BT but not in the BT complex (s25, 2e6, and 2b6) are specific to the binding domain. Thus, under the recombination conditions, BT and HA do not fully reassemble. Some of the binding domain is left uncovered by HA.

Surprisingly, all of the epitopes that were covered in the HA-BT complex were mapped to the binding domain, strongly suggesting that the interactions between HA and BT are mediated by the binding domain. This idea is biologically relevant, since uncomplexed BT is susceptible to trypsin cleavage at the binding domain (22). Furthermore, the trypsinized BT could not bind to brain synaptosomes. When uncomplexed BT was incubated with trypsin, the translocation and catalytic domains showed no sign of proteolysis. Therefore, in the HA-BT complex, the HA may protect the binding domain of BT from proteolytic attack. These observations may guide in developing

antibodies for therapeutic design of neutralizing antibodies against botulism poisoning.

ACKNOWLEDGMENTS

We thank Bill Tepp and Bibhuti DasGupta, Department of Food Microbiology and Toxicology, University of Wisconsin, Madison, for generous supplies of BT and HA-BT complex. Also, we thank Birgitte Andersen for a careful reading of the manuscript.

We gratefully appreciate the support of U.S. Department of Army contracts DAMD17-93-C-3118 and DAMD17-94-C-4034.

REFERENCES

1. **Amersdorfer, P., et al.** Unpublished data.
1a. **Blasl, J., E. R. Chapman, E. Link, T. Binz, S. Yamasaki, P. De Camill, T. C. Südhof, H. Niemann, and R. Jahn.** 1993. Botulinum neurotoxin A selectively cleaves the synaptic protein SNAP-25. Nature **365:**160–163.
2. **Blauestein, R. O., W. J. Germann, A. Finklestein, and B. R. DasGupta.** 1987. The N-terminal half of the heavy chain of botulinum type A neurotoxin forms channels in planar phospholipid bilayers. FEBS Lett. **226:**115–120.
3. **DasGupta, B. R., and D. A. Boroff.** 1968. Separation of toxin and hemagglutinin from crystalline toxin of *Clostridium botulinum* type A by anion exchange chromatography and determination of their dimensions by gel filtration. J. Biol. Chem. **243:**1065–1072.
4. **DasGupta, B. R., D. A. Boroff, and E. Rothstein.** 1966. Chromatographic fractionation of the crystalline toxin of *Clostridium botulinum* type A. Biochem. Biophys. Res. Commun. **22:**750–756.
5. **DasGupta, B. R., and V. Sathyamoorthy.** 1984. Purification and amino acid composition of type A botulinum neurotoxin. Toxicon **22:**415–424.
6. **De Bellis, D., and I. Schwartz.** 1990. Regulated expression of foreign genes fused to lac: control by glucose levels in growth medium. Nucleic Acids Res. **18:**1311.
7. **Ekong, T. A. N., K. McLellan, and D. Sesardic.** 1995. Immunological detection of *Clostridium botulinum* toxin type A in therapeutic preparations. J. Immunol. Methods **180:**181–191.
8. **Fling, S. P., and D. S. Gregerson.** 1986. Peptide and protein molecular weight determination by electrophoresis using a high-molarity Tris buffer system without urea. Anal. Biochem. **155:**83–88.
9. **Hayashi, T., H. McMahon, S. Yamasaki, T. Binz, Y. Hata, T. C. Südhof, and H. Niemann.** 1994. Synaptic vesicle membrane fusion complex: action of clostridial neurotoxins on assembly. EMBO J. **13:**5051–5061.
10. **Hoogenboom, H. R.** 1991. Multisubunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. Nucleic Acids Res. **19:**4133–4137.
11. **Knox, J. N., W. P. Brown, and L. Spero.** 1970. The role of sulfhydryl groups in the activity of type A botulinum toxin. Biochim. Biophys. Acta **214:**350–354.
12. **Kozaki, S., Y. Kamata, T. Nagai, J. Ogasawara, and G. Sakaguchi.** 1986. The use of monoclonal antibodies to analyze the structure of *Clostridium botulinum* type E derivative toxin. Infect. Immun. **52:**786–791.
13. **Kozaki, S., A. Miki, Y. Kamata, T. Nagai, J. Ogasawara, and G. Sakaguchi.** 1989. Immunological characterization of papain-induced fragments of *Clostridium botulinum* type A neurotoxin and interaction of the fragments with brain synaptosomes. Infect. Immun. **57:**2634–2639.
14. **Kozaki, S., S. Nakaue, and Y. Kamata.** 1995. Immunological characterization of the neurotoxin produced by *Clostridium botulinum* type A associated with infant botulism in Japan. Microbiol. Immunol. **39:**767–774.
14a. **Lacy, B.** Unpublished data.
15. **Lamanna, C.** 1948. Hemagglutination by botulinal toxin. Proc. Soc. Exp. Biol. Med. **69:**332–336.
16. **Marks, J. D., H. R. Hoogenboom, T. P. Bonnert, J. McCafferty, A. D. Griffiths, and G. Winter.** 1991. By-passing immunization. Human antibodies from V-gene libraries displayed on phage. J. Mol. Biol. **222:**581–597.
17. **Munro, S., and H. R. Pelham.** 1986. An HSP70-like protein in the ER: identity with the 78 KDa glucose-regulated protein and immunoglobulin heavy chain binding protein. Cell **46:**291–300.
18. **Sathyamoorthy, V., B. R. DasGupta, J. Foley, and R. L. Niece.** 1988. Botulinum neurotoxin type A: cleavage of the heavy and light chains into two halves and their partial sequences. Arch. Biochem. Biophys. **266:**142–151.
19. **Sathyamoorthy, V., B. R. DasGupta, and R. L. Niece.** 1986. Botulinum neurotoxin type A, cleavage and partial sequence of the H chain. Fed. Proc. **45:**1793.
20. **Schier, R., R. F. Balint, A. McCall, G. Apell, J. W. Larrick, and J. D. Marks.** 1996. Identification of functional and structural amino-acid residues by parsimonious mutagenesis. Gene **169:**147–155.
21. **Schier, R., J. Bye, G. Apell, A. McCall, G. P. Adams, M. Malmquist, L. M. Weiner, and J. D. Marks.** 1996. Isolation of high-affinity monomeric human anti-c-erbB-2 single chain Fv using affinity-driven selection. J. Mol. Biol. **255:**28–43.
21a. **Sheets, M.** Unpublished data.
22. **Shone, C. C., P. Hambleton, and J. Melling.** 1985. Inactivation of Clostridium botulinum type A neurotoxin by trypsin and purification of two tryptic fragments. Eur. J. Biochem. **151:**75–82.
23. **Simpson, L. L.** 1980. Kinetic studies on the interaction between botulinum toxin type A and the cholinergic neuromuscular junction. J. Pharmacol. Exp. Ther. **212:**16–21.
24. **Sugii, S., I. Ohishi, and G. Sakaguchi.** 1977. Correlation between oral toxicity and in vitro stability of *Clostridium botulinum* type A and B toxins of different molecular sizes. Infect. Immun. **16:**910–914.
25. **Sugii, S., I. Ohishi, and G. Sakaguchi.** 1977. Intestinal absorption of botulinum toxins of different molecular sizes in rats. Infect. Immun. **17:**491–496.
26. **Sugiyama, H.** 1980. *Clostridium botulinum* neurotoxin. Microbiol. Rev. **44:** 419–448.
27. **Sugiyama, H., I. Ohishi, and B. R. DasGupta.** 1974. Evaluation of type A botulinal toxin assays that use antitoxin to crystalline toxin. Appl. Microbio. **27:**333–336.
28. **Wong, C., et al.** Unpublished data.

Editor: A. O'Brien

JOURNAL OF STRUCTURAL BIOLOGY **120,** 78–84 (1997)
ARTICLE NO. SB973910

Electron Density Projection Map of the Botulinum Neurotoxin 900-kilodalton Complex by Electron Crystallography

Frederick Burkard

Lawrence Berkeley National Laboratory, Berkeley, California 94720

Flora Chen

Graduate Group in Biophysics, University of California, Berkeley, California 94720

Geoffrey M. Kuziemko

Department of Chemistry, University of California, Berkeley, California 94720

and

Raymond C. Stevens[1]

Lawrence Berkeley National Laboratory, Berkeley, California 94720; and Graduate Group in Biophysics and Department of Chemistry, University of California, Berkeley, California 94720

Received July 24, 1997, and in revised form July 31, 1997

The 900-kDa botulinum neurotoxin complex serotype A has been crystallized by the lipid-layer two-dimensional crystallization technique. Based on the binding characteristics of the hemagglutinating portion of the complex, a number of ganglioside/lipid mixtures were tested but only lactosyl ceramide/1-palmityl-2-oleoyl-sn-glycero-3-phosphocholine was found to crystallize the complex. The optimum lipid mixture contained 75 mass % lactosyl ceramide and 25 mass % 1-palmityl-2-oleoyl-sn-glycero-3-phosphocholine. Using protein concentrations from 5 to 500 μg/ml and pH 5 acetate buffer, we have obtained crystals that diffract to better than 15 Å when prepared in negative stain. A projection map with a resolution of 30 Å was calculated with unit cell dimensions of $a = b = 157$ Å and *P*3 symmetry. The complex is triangular in shape with six distinct lobes observed. Additionally, six smaller structures protrude from the triangular core.

[1] To whom correspondence should be addressed at Department of Chemistry, University of California, Berkeley, Berkeley, CA 94720. Fax: (510) 643-9290. E-mail: stevens@adrenaline.berkeley.edu.

INTRODUCTION

Botulinum neurotoxin complex serotype A is a 900-kDa protein produced as one of eight serotypes (A–G) by the anaerobic bacterium *Clostridium botulinum*. Among the most potent biological toxins known, botulinum neurotoxin causes inhibition of synaptic vesicle release at the neuromuscular junction, resulting in flaccid paralysis (Simpson, 1980) and ultimately death. The 900-kDa complex contains a 150-kDa toxic component (BoNTA) and an ~750-kDa component (HA) (DasGupta and Boroff, 1968). Under denaturing and reducing conditions, the complex runs as nine bands on polyacrylamide gels (DasGupta, 1980). The purified 150-kDa neurotoxin runs as a 100-kDa heavy chain and a 50-kDa light chain, which are linked by a disulfide bond under nonreducing conditions. The heavy chain contains the binding and translocation domains (Sugiyama, 1980) which bind to a cell surface receptor and allow the toxin to pass across the membrane, while the light chain contains the catalytic domain which cleaves a protein involved in synaptic vesicle docking (Hayashi *et al.*, 1994). The remaining seven bands of the complex make up the HA component which is nontoxic and has the ability to agglutinate cells. The

1047-8477/97 $25.00

stoichiometry of subunits making up the HA is unclear. The nontoxic components (HA) of the complex are thought to aid the toxin during its travel through the harsh conditions of the gastrointestinal tract.

The HA binds tightly to certain sugars such as *o*-nitrophenyl-*b*-D-galactoside, isopropyl-*b*-D-thiogalactoside, and lactose (DasGupta and Sugiyama, 1977), consequently inhibiting its ability to agglutinate. Furthermore, the complex may be adsorbed onto a sugar affinity column and BoNTA eluted away from the HA by raising the pH from 6.3 to 7.9 (Moberg and Sugiyama, 1978). Several of the sugars recognized by HA appear as head groups on naturally occurring gangliosides, making them ideal candidates as the monolayer used in the two-dimensional lipid-layer crystallization. This method (Uzgiris and Kornberg, 1983) has been used for the crystallization of a number of soluble proteins at the air–water interface and has the advantage of allowing trials with a very small amount of protein over a wide range of buffer conditions. In general, the interaction with the protein is based on either a specific lipid head group such as biotinylated PE with streptavidin (Darst *et al.*, 1991) and GM-1 with cholera toxin (Mosser and Brisson, 1991) or more nonspecific electrostatic interactions such as octadecylamine with RNA polymerase (Darst *et al.*, 1988).

There have been several examples of two-dimensional crystallization of clostridial toxins. Boroff and co-workers (1972) have visualized two-dimensional crystals of the botulinum neurotoxin serotype A toxic and nontoxic components separately using electron microscopy. Botulinum neurotoxin serotype B has been crystallized on lipid layers containing the ganglioside GT1b (Morgan *et al.*, 1989). Schmid *et al.* (1993) visualized the toxic component of serotype B as ordered arrays of channels in phospholipid vesicles and performed a three-dimensional reconstruction. Robinson and co-workers (1988) crystallized tetanus toxin on phospholipid monolayers and performed a three-dimensional reconstruction to 14 Å. Although the botulinum neurotoxin complex readily forms three-dimensional needle-like crystals (Sugiyama *et al.*, 1977), larger crystals are needed for structure determination by X-ray crystallography. To our knowledge, no prior structural information has been published for the 900-kDa complex of botulinum neurotoxin.

After an initial screening of several gangliosides and commonly used lipids, we found that only lactosyl ceramide, a ganglioside with a lactose head group, induced the botulinum neurotoxin complex to crystallize. Subsequent experiments involved optimization of lipid mixtures, protein concentration, and subphase conditions. Two-dimensional crystals grown under optimized conditions were collected, and selected images were processed and used to generate the projection maps presented in this work. Although spots visible to the eye extending to a resolution of 22 Å may be seen in the computed Fourier transform (Fig. 2A) the image processing software is able to identify spots with suitable signal-to-noise ratio to 14 Å resolution (Fig. 2B). Due to the relative scarcity of higher resolution spots and their low amplitudes, we believe that the projection maps that are presented more realistically represent the BoNTA–HA complex to a resolution of 30 Å.

MATERIALS AND METHODS

Botulinum toxin complex was obtained as an ammonium sulfate precipitate and purified as previously described (Chen *et al.*, 1997). The purity of the protein used for crystallization was verified by SDS–PAGE. Botulinum neurotoxin complex was dialyzed into 50 m*M* sodium acetate buffer, pH 5, 100 m*M* sodium chloride and stored at 1.5 mg/ml at 20°C (storage at 4°C causes irreversible sticking to the wall of the container; surprisingly, the complex is stable at 20°C for extended periods of time). Additionally, 10 m*M* citric acid, pH 5, 100 m*M* sodium chloride buffer was also used. All reagents were obtained from Fisher Scientific (Fair Lawn, NJ) unless otherwise noted. Crystallization of the complex was also attempted with 50 or 100 m*M* calcium chloride added to each buffer. Trials with buffer pHs of 4 to 9 were carried out in buffers containing 100 m*M* sodium chloride, 2 m*M* sodium azide. The following buffering agents were used: 10 m*M* citric acid (pH 4), 10 m*M* Bis Tris (pH 6), 10 m*M* Hepes (pH 7 and pH 8), and 10 m*M* CHES (pH 9). 1-Palmityl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) was purchased from Avanti Polar Lipids, Inc. (Alabaster, AL) and stored as a 25 mg/ml stock solution in chloroform at −20°C. Lactosyl ceramide (LC) was purchased from Matreya, Inc. (Pleasant Gap, PA) and stored at 0.5 mg/ml in chloroform at −20°C. The lipid–ganglioside mixtures were 0.5 mg/ml in chloroform with lipid:ganglioside ratios of 100% POPC:0% LC, 95% POPC:5% LC, 90% POPC:10% LC, 80% POPC:20% LC, 70% POPC:30% LC, 60% POPC:40% LC, 50% POPC:50% LC, 35% POPC:65% LC, 25% POPC:75% LC, 15% POPC:85% LC, 0% POPC:100% LC. Crystallization with other gangliosides (Trisialoganglioside GT1b, Monosialogangioside GM1, globotriosyl ceramide, globotetraosyl ceramide GL-4) (Matreya, Inc.) was also attempted. Grids (300 mesh) (Electron Microscopy Sciences, Fort Washington, PA) were coated with carbon and rendered hydrophobic by glow discharge treatment at 0.20 Torr in an atmosphere of hexafluoropropylene (PCR Inc., Gainesville, FL).

Protein was diluted into the appropriate buffer to a concentration between 5 and 500 μg/ml and 20 μl was pipetted into Teflon wells 1 mm deep and 4 mm in diameter. Then 0.5 to 1 μl of lipid–ganglioside solution was placed on top of the protein drop (Darst *et al.*, 1988). Specimens were incubated for 36 to 48 hr at room temperature in a sealed chamber at high humidity to prevent evaporation. Shortly before harvesting, carbon-coated electron microscopy grids were rendered hydrophobic by glow discharge treatment as described above and placed horizontally onto the surface of the protein drop. The grid was removed from the surface of the drop, washed with 50 μl water, and stained with 50 μl 2% (w/v) uranyl acetate (Fluka, Ronkonkoma, NY) for approximately 90 sec.

All experiments were done at room temperature. Preliminary electron microscopy was performed on a Zeiss 10A at 80 kV. Micrographs used for image processing were recorded under low-dose conditions (approximately 10 electrons per $Å^2$) at 40 000× magnification and 2 μm under focus on a Jeol 100B with a field emission gun at 100 kV on Kodak S0-163 film. The magnification

on the Jeol 100B was calibrated with a 2160 lines/mm replica grating and estimated to be 43 900× for the images recorded. Promising micrographs were selected by optical diffractometry. Eight regions on seven micrographs were scanned on a Perkin–Elmer PDS 1010M flat bed microdensitometer controlled by an IBM PC running software developed at LBNL. Images were recorded in arrays of 1000 by 1000 using a 20-μm square aperture. Calculation of the lattice parameters and projection map was done using standard methods (Crowther *et al.*, 1996). Structure factors from individual images were used in the final map only if their signal-to-background ratio was better that 2.5 to 1.

RESULTS

Once it was established that BoNTA–HA would crystallize under a monolayer of LC/POPC, a number of experiments were performed to determine the conditions for optimal size and coherence. In general the crystallization of BoNTA–HA was not especially sensitive to buffer pH, ionic strength, or protein concentration. Protein could be observed adsorbed to the lipid monolayer after 18 to 24 hr of incubation at room temperature and ordered arrays were found most reliably after 36 hr. The lipid composition, however, affected the quality and compactness of the crystalline patches. Lipid mixtures containing from 50 to 100% LC, protein concentrations of 5 to 500 μg/ml, and buffer pHs from 4 to 7 (100 m*M* salt) all yielded crystals that diffracted well. The best conditions were found to be a lipid composition of 75:25 LC to POPC, 7.5 μg/ml BoNTA–HA, and pH 5 in acetate buffer.

Crystals obtained above pH 7 were smaller, were less compact, and diffracted poorly. It is known that the complex dissociates into BoNTA and HA above pH 7.9 (DasGupta and Boroff, 1968). Under the conditions of crystallization, the complex between BoNTA and HA should remain intact. The poor crystal quality seen above pH 7 may be attributed to the high polydispersity observed for the complex at high pH using dynamic light scattering (Chen *et al.*, 1997b).

Figure 1A is representative of the crystalline areas imaged by electron microscopy. As can be seen, the monomeric units are predominantly triangular. Although locally well ordered, there are still disordered regions in the surface, possibly due to imperfect transfer from the air–water interface to the carbon film of the grid. Also noticeable are portions of several lattice rows stained in such a way that they stand out from the rest of the array in Fig. 1A. Closer examination suggests that they are composed of structures different from the "triangles" making up the bulk of the array. Additionally, lattice rows crossing these boundaries typically appear to be shifted approximately ¼ to ⅓ unit cell.

Figure 1B shows both partially formed crystals and what appear to be numerous monomers bound to the monolayer but not yet incorporated into the lattice. As shown by the object circled in this figure, they appear to be composed of a central triangular shape similar in size and shape to the repeating unit seen in the crystal, but surrounded by six structures radiating from the center. Several individual units may be seen in Fig. 1C. Based on measurements on the micrograph, we estimate that these figures have an overall diameter of 200 to 250 Å each. Since the crystalline repeat distance is of the order of 150 Å, it is unclear as to what happens to the radiating structures when the monomer is incorporated into the lattice.

The area marked in Fig. 1D shows a region in which the protein seems to be interacting with the monolayer at two distinct orientations as evidenced by the alternating rows of triangular monomers and ladder-like shapes. This phenomenon gives rise to the striated appearance.

Figure 2A shows a typical computed diffraction pattern resulting from two cycles of correction for lattice distortions (Crowther *et al.*, 1996). The average lattice parameters are $a^* = 1/135 \pm 1/3$ Å^{-1}, $b^* = 1/139 \pm 1/5$ Å^{-1}, and $\gamma^* = 60° \pm 2°$. While there is always some difference in the length of the lattice vectors, we have not identified any features in the diffraction pattern or projection map consistent with the unique assignment of either vector. Based on the unit cell parameters and diffraction data, *P*3 symmetry is assumed at the present resolution. The highest resolution spots easily seen by eye in this pattern are at a resolution of approximately 1/22 Å^{-1} (the −4,7, the 7,−3 and the 3,4 are circled). Figure 2B is a plot of the same data generated by the image processing software. The relative sizes of the circles represent the signal-to-noise ratio of the corresponding diffraction spots. Only spots with a signal-to-noise ratio of 2.5:1 or better are shown. The resolution of the outermost spots is at approximately 1/14 Å^{-1}. The circled spots are the same as in Fig. 2A. Data from seven additional images were processed for lattice distortions, corrected for contrast transfer function phase reversals, combined by vector addition, and weighted on the basis of phase agreement to generate the projection map shown in Fig. 3A. The average phase residual from the data used to generate this map is 18.8° ± 5.0°. The figure shows two by two unit cells with each parallelogram having cell parameters of $a = 155(\pm 3)$ Å, $b = 159(\pm 5)$ Å, and $\gamma = 120°$. Figure 3B shows the same data with *P*3 symmetry imposed; the unit cell parameters are $a = b = 157$ Å and $\gamma = 120°$ with a phase residual of 17.5° ± 4.1°.

Although the results indicate that relatively high-resolution information is present in the data, the amplitudes of the structure factors drop off rapidly at resolutions greater than 1/130 Å^{-1}. This is consistent with the detail seen in the projection maps. The

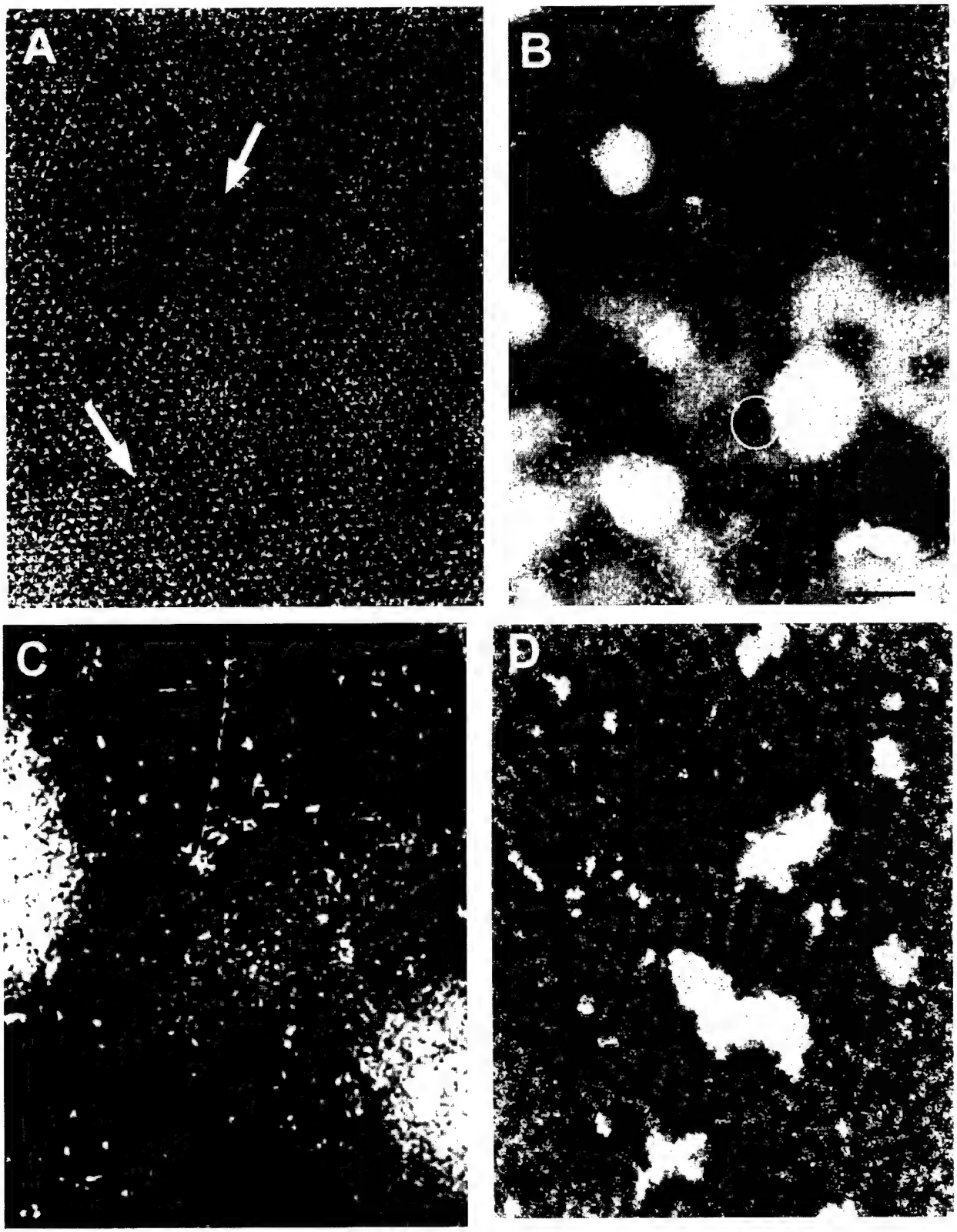


FIG. 1. Electron micrograph of lipid layers transferred to the EM grid and stained with uranyl acetate. BoNTA–HA complex in 100 m*M* sodium chloride, 50 m*M* sodium acetate buffer (pH 5.0) incubated with 3:1 (v/v) lactosyl ceramide: POPC. Bar represents 100 nm. (A) Two-dimensional crystal of botulinum neurotoxin complex. Arrows show lattice dislocations. (B) Monomers of protein adsorbed to monolayer both individually and incorporated into crystal. (C) Enlarged view of monomer unit. Note structures radiating from edges and sides of individual triangular units. (D) Alternating rows of triangular and ladder-like monomers.

sharpness of the visible spots seen in Fig. 2A and the number of reflections with good signal-to-noise ratios apparent in Fig. 2B led us to believe that crystals of the BoNTA–HA complex are locally quite well ordered. Improvements in the size of the crystal, elimination or avoidance of crystal defects, and better transfer from the air–water interface will allow us to improve the resolution of the map.

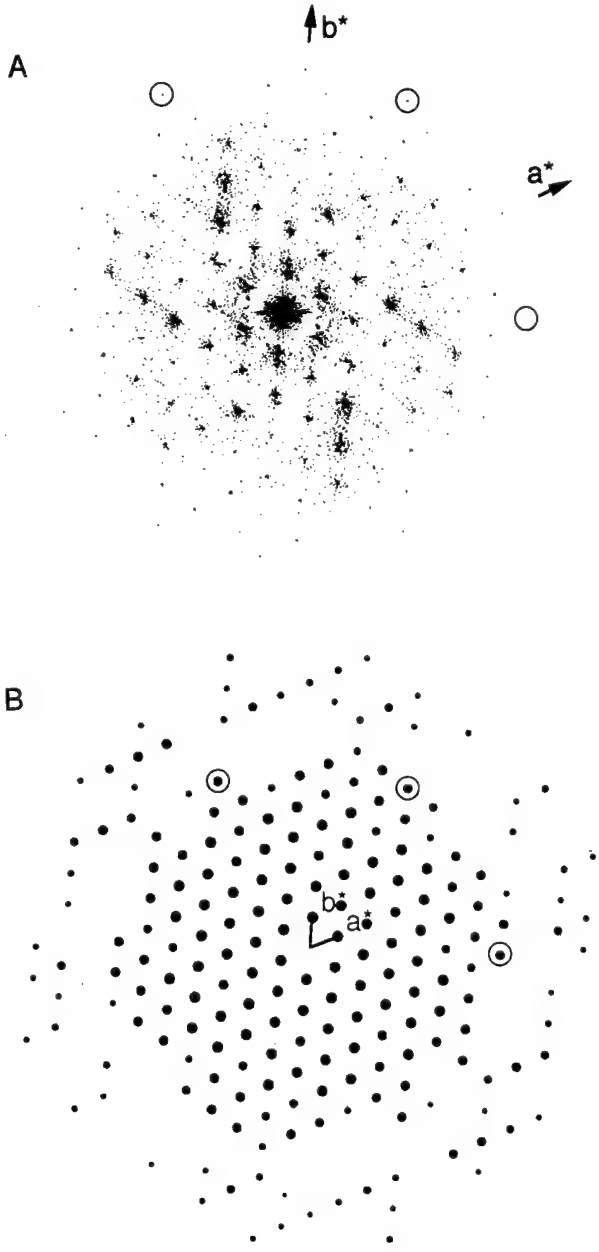


FIG. 2. (A) Computed diffraction pattern corrected for lattice distortions (Henderson *et al.*, 1986). The directions of the reciprocal lattice vectors a^* and b^* are shown. $a^* = 135$ Å^{-1}, $b^* = 139$ Å^{-1}, $\gamma^* = 60°$. The highest resolution reflections apparent to the eye are circled. These are the 2,5, the 7,−3, and the 6,−6 at ~1/22 Å^{-1} resolution. (B) "IQ" plot from same data. Size of filled circle reflects signal-to-noise ratio of spot at lattice position. Spot size ranges from largest (IQ = 1, s:n > 7:1) to smallest (IQ = 4, s:n = 2.5:1). Circled spots are same as in A.

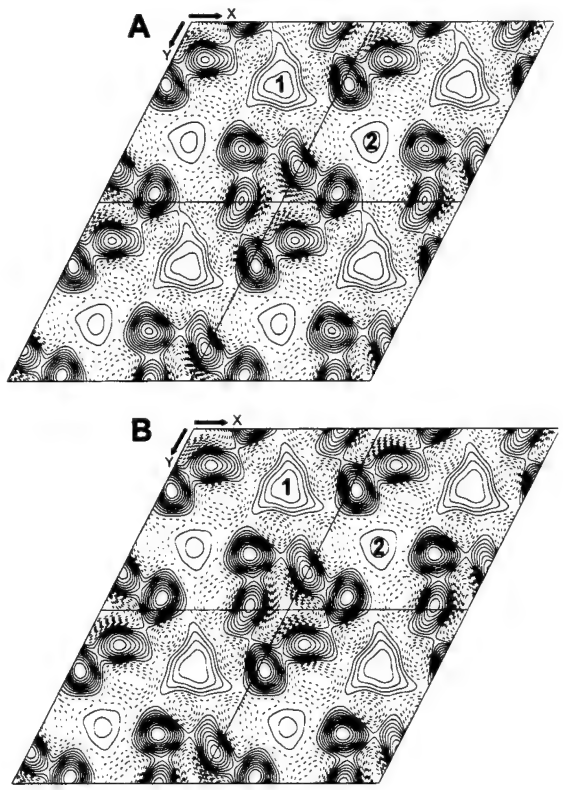


FIG. 3. Projection density maps calculated using noise-filtered Fourier coefficients merged from eight untilted images. Unbroken contours represent densities greater than the mean of the image. Each rhomboid represents one unit cell. (A) Projection map assuming no symmetry constraints (P1). (B) Projection density map calculated assuming *P*3 symmetry.

DISCUSSION

The impetus for structural studies of the 900-kDa complex is that although its components have been well characterized, their stoichiometry and structural organization have not. While an initial investigation of the complex in negative stain yielded no structural information, the lipid-layer technique resulted in the crystals presented in this work. To make sense of the resultant projection map we will try to relate it to what little structural information is available.

The electron micrographs presented in Figs. 1B and 1C show that the structure of a monomeric complex is a central triangular density surrounded by radiating densities. These features are more clearly seen in the projection map of the complex (Figs. 3A and 3B), where they appear as a central triangle, composed of six densities, surrounded by two smaller, triangular densities, labeled "1" and "2." We believe that the less dense structures labeled 1

and 2 in Fig. 3A may be the radiating densities surrounding the incorporated monomers seen in Fig. 1B. Although we know which densities are found within the unit cell, we cannot define the molecular boundary of the complex with certainty. Thus in a single noncrystallized complex, densities 1 and 2 might be adjacent to one another on one side of the central triangle or at opposite ends of the central triangle.

It has previously been shown by sedimentation ultracentrifugation and diffusion constant measurements that the BoNTA–HA has a molecular frictional ratio (f/f_0) of 1.76 (Putnam *et al.*, 1948). From this value it was determined that the complex is a prolate ellipsoid with a molecular weight of 900 kDa. Results obtained from dynamic light scattering indicate that the radius is 110 Å and the molecular weight of the complex is 1200 kDa (Chen *et al.*, 1997b), assuming that the complex is spherical. Since the molecular weight calculated by dynamic light scattering is based on the hydrodynamic radius, and this is calculated on the assumption that the molecule is a sphere, deviations from a spherical shape will overestimate the calculated molecular weight. Based on the discrepancy between calculated molecular weights, it can be inferred that the complex is not spherical but elongated with a major axis of 220 Å.

An approximation of a protein's molecular weight based on its image in projection can be made by assuming the protein to be a sphere with a density of 1.3 g/cm^3 (Cantor and Schimmel, 1980). This approximation works quite well on published maps such as the 133-kDa tetanus toxin (Robinson *et al.*, 1988) and the 117-kDa HIV-1 reverse transcriptase (Kubalec *et al.*, 1994). A 900-kDa spherical protein would have a diameter of 130 Å, fitting nicely into the unit cell but bearing no resemblance to the projection map, which consists of six central densities and two projecting triangular densities, 1 and 2 (Figs. 3A and 3B). Applying this approximation to the six central circular densities, and assuming each individual diameter is 40 Å, gives a total mass of 250 kDa, which falls short of the 900-kDa molecular weight. A better approximation is to consider the six central densities as cylinders instead of spheres. Stretching these densities into cylinders to bring the total mass to 900 kDa requires them to be 100 to 150 Å high. Since the two projecting triangular densities (1 and 2) contribute mass to the complex, then the height of the cylinders should be approximately 100 Å.

We have included both the *P*1 and the *P*3 structures (Figs. 3A and 3B) in this study. Although our projection map shows threefold symmetry, there is no biochemical evidence that BoNTA–HA is trimeric. Additionally, it is implausible that the unit cell contains three 900-kDa protein complexes. Thus, we leave open the possibility that a higher resolution structure may show a reduced degree of symmetry. To reconcile the biochemical data with the electron microscopy, we propose the following model: The 900-kDa BoNTA–HA complex is composed of six cylinders approximately 40 Å in diameter by 100 Å high. These are arranged in a triangular pattern 130 Å on each side and hollow in the middle. Six smaller structures project radially 110 Å from the center of the triangle (yielding a 220-Å diameter in the plane of the projection map). Verification of this structure will require the generation of a higher resolution map both in projection and as a three-dimensional reconstruction.

We gratefully acknowledge Drs. Kenneth Downing and Robert Glaeser for many helpful discussions on electron microscopy and image processing. We also thank Dr. Bibhuti DasGupta and Bill Tepp for their generosity in supplying BoNTA–HA. This work was supported by the National Institutes of Health (Grants GM51487 and GM36884) and the Office of Health and Environmental Research, U.S. Department of Energy (under Contract DE-AC03-76F00098).

REFERENCES

Amos, L. A., Henderson, R., and Unwin, P. N. T. (1982) Three-dimensional structure determination by electron microscopy of two dimensional crystals, *Prog. Biophys. Mol. Biol.* **39,** 183–231.

Boroff, D. A., Nyberg, S., and Hoglund, S. (1972) Electron microscopy of the toxin and hemagglutinin of type A *Clostridium botulinum, Infect. Immun.* **6**(6), 1003–1007.

Cantor, C. R., and Schimmel, P. R. (1980) Biophysical Chemistry, Part 1, The Conformation of Biological Macromolecules, pp. 115–116, Freeman, New York.

Chen, F., Kuziemko, G. M., Amersdorfer, P., Wong, C., Marks, J. D., and Stevens, R. C. (1997) Antibody mapping to domains of botulinum neurotoxin serotype A in the complexed and uncomplexed forms, *Infect. Immun.* **65**(5), 1626–1630.

Chen, F., Kuziemko, G. M., and Stevens, R. C. Submitted for publication.

Crowther, R. A., Henderson, R., and Smith, J. M. (1996) MRC image processing programs, *J. Struct. Biol.* **116,** 9–16.

Darst, S. A., Ahlers, M., Meller, P. H., Kubalek, E. W., Blankenburg, R., Ribi, H. O., Ringsdorf, H., and Kornberg, R. D. (1991) Two-dimensional crystals of streptavidin on biotinylated lipid layers and their interaction with biotinylated macromolecules, *Biophys. J.* **59,** 387–396.

Darst, S. A., Ribi, H. O., Pierce, D. W., and Kornberg, R. D. (1988) Two-dimensional crystals of *Escherichia coli* RNA polymerase holoenzyme on positively charged lipid layers, *J. Mol. Biol.* **203,** 269–273.

DasGupta, B. R. (1980) Electrophoretic analysis of clostridium botulinum types A and B hemagglutinins, *Can. J. Microbiol.* **26,** 992–997.

DasGupta, B. R., and Boroff, D. A. (1968) Separation of toxin and hemagglutinin from crystalline toxin of *Clostridium botulinum* type A by anion exchange chromatography and determination of their dimensions by gel filtration, *J. Biol. Chem.* **243**(5), 1065–1072.

DasGupta, B. R., and Sugiyama, H. (1977) Inhibition of Clos-

tridium botulinum types A and B by sugars, *Can. J. Microbiol.* **23,** 1257–1260.

Hayashi, T., McMahon, H., Yamasaki, S., Binz, T., Hata, Y., Südhof, T. C., and Niemann, H. (1994) Synaptic vesicle membrane fusion complex: Action of clostridal neurotixins on assembly, *EMBO J.* **13,** 5051–5061.

Henderson, R., Baldwin, J. M., Downing, K. H., Lepault, J., and Zemlin, F. (1986) Structure of purple membrane from *Halobacterium halobium:* Recording, measurement and evaluation of electron micrographs at 3.5 Å resolution, *Ultramicroscopy* **19,** 147–178.

Kubalek, E. W., LeGrice, S. F. J., and Brown, P. O. (1994) Two-dimensional crystallization of histidine-tagged, HIV-1 reverse transcriptase promoted by a novel nickel-chelating lipid, *J. Struct. Biol.* **113,** 117–123.

Moberg, L. J., and Sugiyama, H. (1978) Affinity chromatography purification of type A botulinum neurotoxin from crystalline toxic complex, *Appl. Environ. Microbiol.* **35,** 878–880.

Morgan, D. G., DasGupta, B. R., Stubbs, B., and Robinson, J. P. (1989) Two dimensional crystals of botulinum toxin, serotype B, Proceedings of the Annual Meeting of the Electron Microscopy Society of America, pp. 1034–1035.

Mosser, G., and Brisson, A. (1991) Conditions of two-dimensional crystallization of cholera toxin B-subunit on lipid films containing ganglioside GM1, *J. Struct. Biol.* **106,** 191–198.

Putnam, F. W., Lamanna, C., and Sharp, D. G. (1948) Physicochemical properties of crystalline *Clostridium botulinum* type A toxin, *J. Biol. Chem.* **176,** 401–412.

Robinson, J. P., Schmid, M. F., Morgan, D. G., and Chiu, W. (1988) Three-dimensional structural analysis of tetanus toxin by electron crystallography, *J. Mol. Biol.* **200,** 367–375.

Simpson, L. L. (1980) Kinetic studies on the interaction between botulinum toxin type A and the cholinergic neuromuscular junction, *J. Pharmacol. Exp.* **212,** 16–21.

Schmid, M. F., Robinson, J. P., and DasGupta, B. R. (1993) Direct visualization of botulinum neurotoxin-induced channels in phospholipid vesicles, *Nature* **364,** 827–830.

Sugiyama, H. (1980) *Clostridium botulinum* neurotoxin, *Microbiol. Rev.* **44,** 419–448.

Sugiyama, H., Moberg, L. J., and Messer, S. L. (1977) Improved procedure for crystallization of *Clostridium botulinum* type A toxic complexes, *Appl. Environ. Microbiol.* **33**(4), 963–966.

Uzgiris, E. E., and Kornberg, R. D. (1983) Two-dimensional crystallization technique for imaging macromolecules, with application to antigen–antibody–complement complexes, *Nature* **301,** 125–129.

PROTEIN EXPRESSION AND PURIFICATION 11, 195–200 (1997)
ARTICLE NO. PT970772

Recombinant Expression and Purification of the Botulinum Neurotoxin Type A Translocation Domain

D. Borden Lacy and Raymond C. Stevens[1]
Department of Chemistry, University of California, Berkeley, California 94720

Received April 29, 1997, and in revised form June 9, 1997

Botulinum neurotoxin type A in its fully activated form exists as a dichain protein consisting of a 50-kDa light chain and a 100-kDa heavy chain linked by a disulfide bond (B. R. DasGupta and H. Sugiyama, *Biochem. Biophys. Res. Commun.* 48, 108–112, 1972). The protein can be further subdivided into three functional domains: a catalytic domain corresponding to the light chain, a translocation domain associated with the N-terminal half of the heavy chain, and a binding domain as the C-terminal half. To facilitate further structural and functional studies on the mechanism of toxin translocation, we report here the recombinant *Escherichia coli* expression and purification of the isolated translocation domain with a yield of 1 mg pure protein per 1 g cell paste. Circular dichroism, enzyme-linked immunosorbent assays, and preliminary crystallization experiments verify proper protein folding. This reagent should serve as a key tool in elucidating the mechanism of translocation and in determining how the catalytic domain, a large 50-kDa metalloprotease, is delivered to the cytosol.

Botulinum neurotoxin serotype A (BoNT/A)[2] is one of seven antigenically distinct proteins produced from strains of *Clostridium botulinum.* These serotypes (designated types A–G) are the causative agents of botulism, a potentially fatal condition of neuromuscular paralysis (2). The seven types of BoNT along with tetanus toxin comprise the clostridial toxin family. Comparison of the clostridial toxins to the better-characterized diphtheria toxin (DT) shows some important similarities. Both the clostridial toxins and DT are synthesized as a single chain and then posttranslationally nicked to form a dichain molecule held together by a disulfide bond. Their toxicity is believed to be a result of three functional domains (Fig. 1a shows BoNT/A) acting in a three-step mechanism to bind, enter, and catalyze a specific reaction within the cell. The idea that BoNT shares this three-step model of toxicity was first proposed by Simpson (3) and has since been supplemented with further experimental evidence and analogies to DT. The first step involves a binding interaction between the 50-kDa binding domain and a cell surface receptor on the presynaptic membrane (4). BoNT is thought to be internalized into an endosome via receptor-mediated endocytosis. The acidic pH of the endosome is believed to then cause a conformational change [best characterized for DT (5, 6)] in the 50-kDa translocation domain which allows it to interact with the endosomal membrane and form a channel (7, 8). The 50-kDa catalytic domain is translocated across the membrane, released from its disulfide linkage, and left free in the cytosol. The catalytic domain of each BoNT serotype is a zinc-dependent endopeptidase which targets specific sequences of proteins involved in vesicle docking and membrane fusion. BoNT/A's cleavage of a peptide bond in SNAP-25 (synaptosomal-associated protein of 25-kDa) (9) results in inhibition of acetylcholine release from the axon and, ultimately, paralysis (10).

Of particular interest is the proposed role of the BoNT translocation domain in the delivery of the catalytic domain to the cytosol: specifically, (i) the nature of the conformational change at acidic pH, (ii) the molecular mechanism of insertion, and (iii) the role the translocation domain plays—active or passive—in getting the 50-kDa catalytic domain into the cytosol.

[1] To whom correspondence should be addressed. Fax: (510) 643-9290.

[2] Abbreviations used: BoNT/A, botulinum neurotoxin type A; BoNT, botulinum neurotoxin; DT, diphtheria toxin; Tris, tris(hydroxymethyl)aminomethane; ECL, enhanced chemiluminescence; IPTG, isopropyl β-D-thiogalactopyranoside; PCR, polymerase chain reaction; EDTA, ethylenediaminetetraacetic acid; BoNT-trans, botulinum neurotoxin translocation domain; CD, circular dichroism; ELISA, enzyme-linked immunosorbent assay; PBS, phosphate-buffered saline; scFv, single-chain variable fragment; Fc, constant fragment; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; PE, *Pseudomonas* exotoxin A.

1046-5928/97 $25.00

A great deal of insight into the phenomenon of translocation has come from studies done on diphtheria toxin, particularly since the elucidation of the atomic coordinates by X-ray crystallography (11). Using biochemical studies, researchers have used this structural information to assess the roles that both secondary structural elements and individual amino acids play in the translocation event (12–16).

In contrast, very little information about translocation is available for the clostridial toxins. While comparisons between toxin families have been useful in the initial understanding of BoNT, significant differences exist. Most notably, DT is a 58-kDa protein while BoNT is 150 kDa. The enormous size of BoNT relative to DT and all other toxins makes it uniquely interesting, especially in the context of cellular transport. Sequence comparisons indicate that when aligned to BoNT/A, the other serotypes and tetanus toxin have 55 to 66% homology. The translocation domains share a slightly higher homology of 60 to 70%. We, therefore, have pursued BoNT/A as a model system for studying translocation of the clostridial toxins.

As a first step in this investigation, the isolated translocation domain of BoNT/A has been cloned into a recombinant *Escherichia coli* overexpression system. This is the first report of translocation domain expression for botulinum neurotoxin along with initial biophysical characterization. Constructs were prepared both with and without cysteines and with silent mutations to improve homogeneity. Given the partial membrane–protein character of the domain, a detailed description of the purification is presented as well as preliminary structural experiments.

EXPERIMENTAL PROCEDURES

Materials. All buffer components and reagents were purchased from Sigma unless otherwise specified: Tris (Fisher Scientific), enhanced chemiluminescence (ECL) reagents (Amersham), DNAse (Boehringer-Mannheim), silver-staining solutions (Bio-Rad), and isopropyl β-D-thiogalactopyranoside (IPTG) (Diagnostic Chemicals, Inc.). Restriction endonucleases were obtained from New England Biolabs, Inc. and used as directed by the supplier. T4 DNA ligase, T4 polymerase, and T4 polynucleotide kinase were obtained from Pharmacia Biotech, Inc. Additionally, Pharmacia supplied the chelating Sepharose fast flow resin and the reagents for fluorescence dideoxy sequencing that accompany their automated laser fluorescent (A.L.F.) DNA sequencer. The oligonucleotides for the polymerase chain reaction (PCR) were purchased from Genset, while the oligonucleotides for site-directed mutagenesis and sequencing were synthesized in-house on an Applied Biosystems 392 RNA/DNA synthesizer. Fluorescent primers were prepared with FluorePrime, a fluorescein amidite, also from Pharmacia.

Bacterial strains. *E. coli* strain DH5αF′ was used in the cloning procedures, *E. coli* strain CJ236 was used for mutagenesis, and *E. coli* strain BL21.DE3.pLysS was used for protein overexpression.

Construction of pET-23a/BoNT/A-trans. Plasmid DNA used for cloning and sequencing was prepared through either the alkaline lysis method (17) or with a kit supplied by QIAGEN. DNA encoding the BoNT/A catalytic and translocation domains was isolated from genomic DNA (18) by PCR and cloned into a pUC119 vector with a C-terminal myc epitope tag (GAACAAAAACTCATCTCAGAAGAGGATCTGAATGGGGCCGCA). This construct was prepared as a linear strand template for PCR to amplify just the translocation domain. Oligonucleotides were designed against the translocation domain and C-terminal myc epitope tag to include two unique restriction sites. The forward primer introduced a *Nde*I site at the 5′ end of the gene (GAAAAAGGGAAGGGACATATGGCATTAAATGATTTATGTATCAAAGTTAATAATTGG) and the reverse primer (GAAAGAAAGAGGGGGCTCGAGTGCGGCCCCATTCAGATCCTC) extended from the myc epitope tag to include a *Xho*I site. PCR products were gel-purified and extracted using the QIAGEN gel extraction kit. Following resuspension in TE (10 mM Tris, 1 mM EDTA, pH 8.0), the DNA was ligated into a pGEMT vector (Promega) by A/T overlaps and sequenced to check for correct PCR amplification. The DNA was then digested with *Nde*I and *Xho*I and ligated into the pET-23a vector (Novagen). The pET-23a *Xho*I site is followed by six histidine residues and the stop codon.

Site-directed mutagenesis. The pET-23a/BoNT/A-trans plasmid was transformed into a CJ236 dut$^-$ ung$^-$ strain of *E. coli.* Site-directed mutagenesis was performed using standard protocols for manipulating single-stranded uracil-containing DNA (17). The oligo for the silent mutations (GCCTAAAAACATAGCCGCTTCGGTCGCTTTATTAACTTTCTTTAC) took advantage of the degeneracy of the genetic code to eliminate the internal Shine–Dalgarno sequence. Two additional oligos were designed to replace the cysteines with alanines: (CAAGTCCCAATTGTTAACTTTGATGGCCAGATCATTTAATGCC) annealed around C454 and (CCATAAGGGATCATAGAGTTCATGAGATAAGAAACGCTAGCTTGATTCAAAAA) for C791. These two oligos were annealed and extended simultaneously to produce the double mutant.

Expression and purification of the translocation domain. Expression of the BoNT/A-trans domain was under the control of the T7 polymerase promoter in pET-23a. Initially, expression was done in 1-liter shaker flasks containing LB/carb/CAP [Luria-Bertani medium (17) containing 50 μg/mL carbenicillin and 34 μg/mL chloramphenicol]. The medium was inoculated with a single colony of BL21.DE3.pLysS carrying the

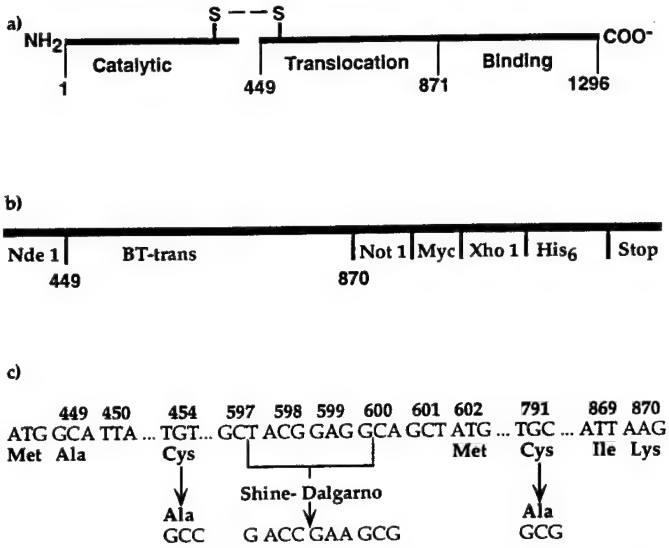


FIG. 1. BoNT/A-translocation domain constructs in pET23a. (a) The three functional domains of BoNT/A. (b) The original PCR product was cloned into pET-23a using *Nde*I and *Xho*I. (c) Site-directed mutagenesis was used to convert cysteines to alanines and to introduce silent mutations in the internal Shine–Dalgarno sequence.

desired plasmid, incubated with shaking at 37°C, and, at an OD = 0.6, induced with 1 mM IPTG. In later experiments, the cells were grown in a 200-liter fermentor with all reagents scaled up proportionally. Starting with 60 g of cell paste, the bacterial pellet was resuspended in 30 mL of binding buffer (500 mM NaCl, 5 mM imidazole, 0.3 mM methionine, and 20 mM Tris, pH 7.9). Protease inhibitors were added to final concentrations of 0.12 mg/mL phenylmethylsulfonyl fluoride (PMSF), 6.7 μg/mL pepstatin A, and 6.7 μg/mL leupeptin. Additionally, 200 μL of 1 M $MgCl_2$ and small amounts of DNAse were added. Cells were lysed using a French press and the soluble and insoluble portions were separated by centrifugation. The pellet was washed with 30 mL binding buffer in the presence of 1 mL 0.5 M EDTA and recentrifuged prior to resuspension in a total volume of 30 mL binding buffer in 1% *n*-lauroyl sarcosine. Following centrifugation, the supernatant was filtered and further purified using metal affinity chromatography. A column packed with metal chelating resin was charged with either 50 mM $NiSO_4$ or $CuSO_4$. The column was pre-equilibrated with the binding buffer plus 1% sarcosyl solution be-

TABLE 1

Purification Summary

Purification step	Volume (mL)	Total protein (mg)
Detergent solubilization	30	ND
Metal affinity column	7.5	60

Note. The data were obtained from purification of the BoNT/A translocation domain from 60 g of cell paste as described under Experimental Procedures, using 10 mL of Pharmacia metal chelating resin.

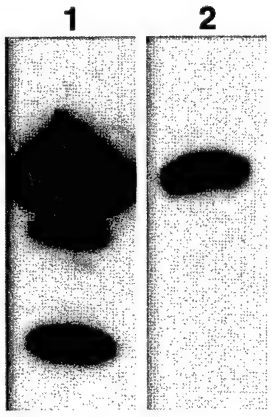


FIG. 2. Western blotting to confirm protein identity. SDS–PAGE was followed by transfer to nitrocellulose, sequential incubation with anti-Myc and anti-mouse horseradish peroxidase antibodies, and detection by ECL. Lane 1 contains 30 μg of protein and shows the translocation domain with lower molecular weight proteolysis and cotranslational contaminants. Lane 2 contains 2 μg of protein and shows just the BoNT/A translocation domain after steps had been taken to avoid the contaminants.

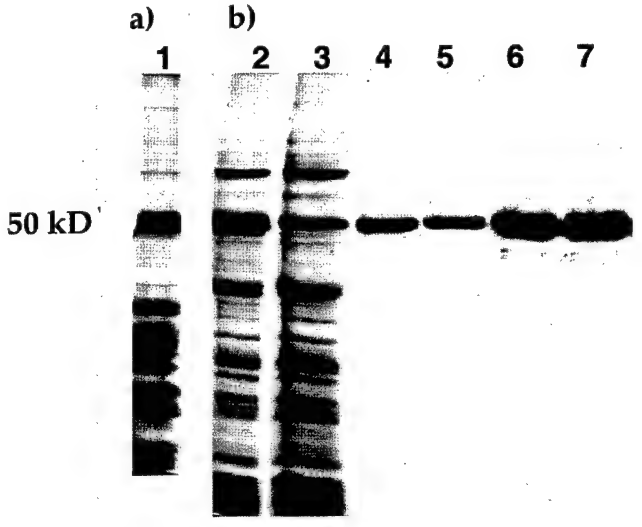


FIG. 3. SDS–PAGE to assess expression and purity of BoNT/A-trans. Electrophoresis was performed on a 10% acrylamide gel under reducing conditions with 30 μg of protein per lane. Protein bands were visualized by silver staining. Lanes 1 through 6 show the purification of the BoNT/A-trans (with all mutations incorporated) by metal affinity chromatography. Lane 1 shows what was loaded on the column, lane 2 represents the flowthrough, and lanes 3–6 show the purified protein after elution with 500 mM imidazole.

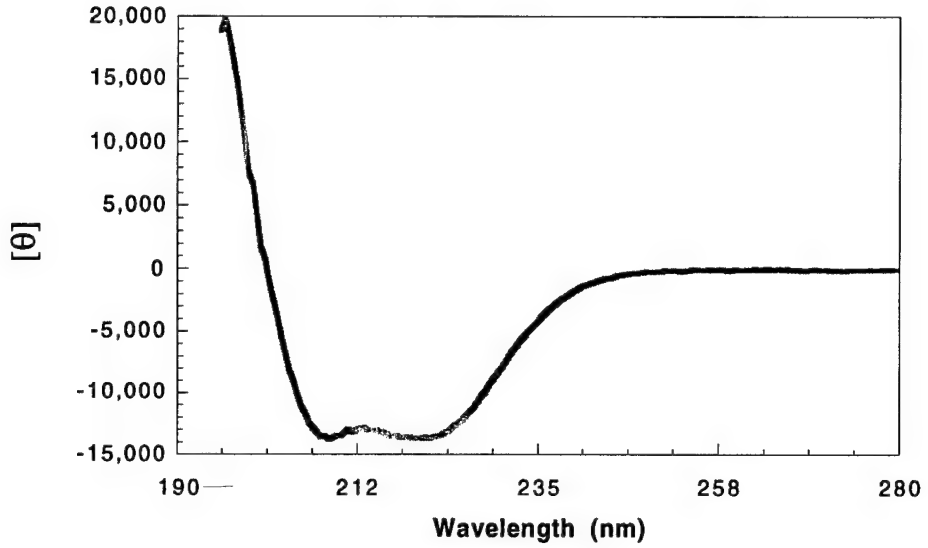


FIG. 4. Far-UV CD spectrum of the BoNT/A-translocation domain. A wavelength scan was taken at 25°C using a 1-cm path-length strain-free cuvette (BoNT/A-trans at 30 μg/mL, 5 mM potassium phosphate, pH 7.0, 5 mM KCl).

fore loading the protein. Following an extended wash step, the protein was eluted off the column by increasing the imidazole concentration to 500 mM.

Circular dichroism (CD). All CD spectra were obtained on an Aviv 62DS spectropolarimeter with a Peltier temperature-controlled sample holder and a 1-cm pathlength cuvette. Potassium phosphate buffers were prepared at pH 3, 6, 7, and 8. Potassium acetate was used as the buffer at pH 4 and 5.

Enzyme-linked immunosorbent assay (ELISA). Microtiter plates (Falcon 3912) were incubated with 50 μL of the translocation domain at 10 μg/mL in PBS (25 mM NaH_2PO_4, 125 mM NaCl, pH 7.0) at 4°C overnight. After washing once with PBS, wells were incubated with 50 μL bacterial supernatant containing native E-tagged single-chain variable fragment (scFv). Antibodies were detected using an anti-E-tag antibody (1 mg/mL) (Pharmacia), followed by peroxidase-conjugated anti-mouse Fc antibody (Sigma), and 2,2′-azinobis(3-ethylbenzthiazoline sulfonic acid) as substrate as described (19).

Crystal growth. Initial crystals were found from an incomplete factorial grid (20) using Crystal Screens I and II (Hampton Research). Crystals were grown at 4 and 25°C by the hanging drop vapor diffusion method from a starting protein concentration of 7.8 mg/mL.

Sequence homology. Sequence homology between the clostridial toxins was determined by the BLAST algorithm (21) to align sequences in the SwissProt database.

RESULTS AND DISCUSSION

Construction of a BoNT/A-trans overexpression system and related mutations. Primers were designed to isolate the translocation domain as residues 449–870 (Fig. 1a) as determined in the complete amino acid sequence of the holotoxin (22). This construct's sequence differs from the target sequence by an N-terminal methionine to start translation, a C-terminal Myc epitope tag to facilitate detection by Western blot analysis, and a C-terminal polyhistidine tag to enable purification by metal affinity chromatography (Fig. 1b). The initial construct contains two cysteines which do not form a disulfide bond in the native protein since the N-terminal cysteine is known to form the disulfide bridge with the catalytic domain. While the initial cysteine-containing construct was shown to overexpress, unwanted inter- or intramolecular disulfides were prevented by replacing the two cysteines with alanines. This construct, BoNT/A-trans (C454A, C791A), was prepared by site-directed mutagenesis, verified by sequencing, and shown to express in similar yields to the original construct. The protein identity was verified by Western blot analysis (Fig. 2, lane 1). Here, the upper 50-kDa band corresponds to the translocation domain. The lower band at 40 kDa was a result of proteolytic cleavage and was removed through the incorporation of protease inhibitors in the purification. Interestingly, the lower band at 30 kDa was a result of an internal methionine codon located six bases downstream from a Shine–Dalgarno sequence in which only one of eight bases differed from the canonical sequence (Fig. 1c). This cotranslational contaminant was eliminated by introducing silent mutations in the internal Shine–Dalgarno sequence by site-directed mutagenesis. Specifically, the codons for Ala 597, Thr 598, Glu 599, and Ala 600 were changed to degenerate alternatives. These changes were again, verified by sequencing, and

new Western blot analysis showed the complete elimination of this contaminant (Fig 2, lane 2), thereby dramatically improving the homogeneity of the sample.

Purification of the BoNT/A-translocation domain. The protein was expressed to the insoluble portion of the cell. After experimenting with a number of salts, denaturants, and detergents, 1% *n*-lauroyl sarcosine, a zwitterionic detergent, was found to be the most effective means of solubilizing the protein. The resolubilized protein was then further purified by metal affinity chromatography (Table 1). The sample (Fig. 3, lanes 3–6) is believed to be greater than 95% pure as assessed by Coomassie and silver stain gel electrophoresis with serial dilutions of protein. This expression and purification protocol yields 1 mg pure protein for every 1 g of cell paste. The translocation domain is folded and stable in solution as shown by CD, ELISA, and crystallization experiments.

CD spectropolarimetry. Since the recombinant protein was isolated from the insoluble fraction, it is especially important to determine if the protein is folded. Far-UV CD spectropolarimetry of the recombinant protein showed a spectrum indicative of secondary structure as evidenced by the minima at 208 and 222 nm (Fig. 4). Similar wavelength scans were taken in 1-unit pH increments from 3 to 8. Two different buffering agents were required to achieve the optimal pK_a's. It was interesting to observe that the spectra at pH 3, 6, 7, and 8 were identical. Unfortunately, spectra could not be attained at pH 4 and 5, as the protein was insoluble under these conditions. This 4–5 pH range is also the pH range of the endosome when the putative conformational change occurs. The calculated p*I* of the protein domain is 3.6. While it is not surprising that the protein would become more insoluble as it rearranges itself to interact with the membrane, further efforts will be made to find conditions where these changes can be monitored by biophysical techniques.

ELISA. Further evidence in support of a properly folded structure comes from work in mapping antibodies binding the holotoxin to the individual domains (23). Of the 40 scFv's that mapped to the holotoxin, three mapped exclusively to the translocation domain. ELISAs on these three antibodies were performed on protein purified as described above and on the same protein after heat denaturation. The denatured protein showed no signal by ELISA, possibly indicating that the antibodies recognize a conformational epitope as opposed to a linear one.

Preliminary crystallization results. The final piece of evidence in favor of structured protein was that the protein has been shown to crystallize in the presence of 100 mM magnesium formate, 100 mM Tris, pH 8.5, and 1 mM sodium azide (Fig. 5). While not yet of diffraction quality, the ability to grow crystals indicates a specific molecular packing arrangement unlikely for denatured proteins. Crystals were confirmed to be protein by staining with methylene blue and by analysis of crystals by SDS–PAGE.

FIG. 5. Preliminary crystallization of the BoNT/A-translocation domain. The protein was concentrated to 7.8 mg/mL and crystallized in the presence of 100 mM magnesium formate. Crystals were grown at 4°C in 2 days by the hanging drop method.

CONCLUSION

The BoNT/A translocation domain is part of a soluble holoprotein that does not become activated until it enters the endosome. Once inside the endosome, the low pH is believed to induce conformational changes in the translocation domain that allow it to merge with the endosomal cell membrane and carry the 50-kDa catalytic domain to the cytosol. Once in the cytosol, the protease is able to perform the final function of the toxin, the inhibition of synaptic vesicle fusion resulting in nerve paralysis. Efforts to understand the basic mechanism of toxin translocation events have been undertaken in a number of systems, including DT (12–

16), *Pseudomonas* exotoxin A (PE) (24), ricin (25), and shiga toxin (26). However, none of these other toxins carry such a large 50-kDa protein to the cytosol. The ability to produce the translocation domain in suitable amounts provides the opportunity to study the translocation event in BoNT/A. Since botulinum neurotoxin is able to carry such a large protein to the cytosol, the translocation domain could possibly be harnessed to deliver other large proteins to their targets within the cytosol, analogous to immunotoxin and protein design strategies pursued previously with DT and PE (27–29). A further understanding of the translocation mechanism will greatly aid in this endeavor.

ACKNOWLEDGMENTS

This work was supported in part by the U. S. Army Medical Research and Development Command (DAMD17-93-C-3118), a NSF predoctoral fellowship (B.L.), and a NSF Young Investigator Award (R.C.S.). We thank Carl Chen for the initial genomic DNA cloning and Flora Chen for performing the ELISA experiments.

REFERENCES

1. DasGupta, B. R., and Sugiyama, H. (1972) A common subunit structure in *Clostridium botulinum* type A, B and E toxins. *Biochem. Biophys. Res. Commun.* **48,** 108–112.
2. Sugiyama, H. (1980) *Clostridium botulinum* neurotoxin. *Microbiol. Rev.* **44,** 419–448.
3. Simpson, L. L. (1980) Kinetic studies on the interaction between botulinum toxin type A and the cholinergic neuromuscular junction. *J. Pharmacol. Exp. Ther.* **212,** 16–21.
4. Dolly, J. O., Black, J., Williams, R. S., and Melling, J. (1984) Acceptors for botulinum neurotoxin reside on motor nerve terminals and mediate its internalization. *Nature* **307,** 457–460.
5. Blewitt, M. G., Chung, L. A., and London, E. (1985) Effect of pH on the conformation of diphtheria toxin and its implications for membrane penetration. *Biochemistry* **24,** 5458–5464.
6. Dumont, M. E., and Richards, F. M. (1988) The pH-dependent conformational change of diphtheria toxin. *J. Biol. Chem.* **263,** 2087–2097.
7. Hoch, D. H., Romero-Mira, M., Ehrlich, B. E., Finkelstein, A., DasGupta, B. R., and Simpson, L. L. (1985) Channels formed by botulinum, tetanus, and diphtheria toxins in planar lipid bilayers: Relevance to translocation of proteins across membranes. *Proc. Natl. Acad. Sci. USA* **82,** 1692–1696.
8. Schmid, M. F., Robinson, J. P., and DasGupta, B. R. (1993) Direct visualization of botulinum neurotoxin-induced channels in phospholipid vesicles. *Nature* **364,** 827–830.
9. Blasi, J., Chapman, E. R., Link, E., Binz, T., Yamasaki, S., De Camilli, P., Südhof, T. C., Niemann, H., and Janh, R. (1993) Botulinum neurotoxin A selectively cleaves the synaptic protein SNAP-25. *Nature* **365,** 160–163.
10. Burgen, A. S. V., Dickens, F., and Zalman, L. J. (1949) The action of botulinum toxin on the neuromuscular junction. *J. Physiol. (London)* **109,** 10–24.
11. Choe, S., Bennett, M. J., Fujii, G., Curmi, P. M. G., Kantardjieff, K. A., Collier, R. J., and Eisenberg, D. (1992) The crystal structure of diphtheria toxin. *Nature* **357,** 216–222.
12. Johnson, V. G., Nicholls, P. J., Habig, W. H., and Youle, R. J., (1993) The role of Proline 345 in diphtheria toxin translocation. *J. Biol. Chem.* **268,** 3514–3519.
13. Madshus, I. H. (1994) The N-terminal α-helix of fragment B of diphtheria toxin promotes translocation of fragment A into the cytoplasm of eukaryotic cells. *J. Biol. Chem.* **269,** 17723–17729.
14. vanderSpek, J., Cassidy, D., Genbauffe, F., Huynh, P. D., and Murphy, J. R. (1994) An intact transmembrane helix 9 is essential for the efficient delivery of the diphtheria toxin catalytic domain to the cytosol of target cells. *J. Biol. Chem.* **269,** 21455–21459.
15. Silverman, J. A., Mindell, J. A., Finkelstein, A., Shen, W. H., and Collier, R. J. (1994) Mutational analysis of the helical hairpin region of diphtheria toxin transmembrane domain. *J. Biol. Chem.* **269,** 22524–22532.
16. Zhan, H., Oh, K. J., Shin, Y.-K., Hubbell, W. L., and Collier, R. J. (1995) Interaction of the isolated transmembrane domain of diphtheria toxin with membranes. *Biochemistry* **34,** 4856–4863.
17. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
18. Betley, M. J., Somers, E., and DasGupta, B. R. (1989) Characterization of botulinum type A neurotoxin gene: Delineation of the N-terminal encoding region. *Biochem. Biophys. Res. Commun.* **162,** 1388–1395.
19. Schier, R., Bye, J., Apell, G., McCall, A., Adams, G. P., Malmquist, M., Weiner, L. M., and Marks, J. D. (1996) Isolation of high-affinity monomeric human anti-c-erbB-2 single chain Fv using affinity-driven selection. *J. Mol. Biol.* **255,** 28–43.
20. Jancarik, J., and Kim, S. H. (1991) Sparse matrix sampling—A screening method for crystallization of proteins. *J. Appl. Crystallogr.* **24,** 409–411.
21. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215,** 403–410.
22. Binz, T., Kurazono, H., Wille, M., Frevert, J., Wernars, K., and Niemann, H. (1990) The complete sequence of botulinum neurotoxin type A and comparison with other clostridial neurotoxins. *J. Biol. Chem.* **265,** 9153–9158.
23. Chen, F., Kuziemko, G. M., Amersdorfer, P., Wong, C., Marks, J. D., and Stevens, R. C. (1997) Antibody mapping to domains of botulinum neurotoxin serotype A in the complexed and uncomplexed forms. *Infect. Immun.* **65,** 1626–1630.
24. Theuer, C. P., Buchner, J., FitzGerald, D., and Pastan, I. (1993) The N-terminal region of the 37-kDa translocated fragment of *Pseudomonas* exotoxin A aborts translocation by promoting its own export after microsomal membrane insertion. *Proc. Natl. Acad. Sci. USA* **90,** 7774–7778.
25. Beaumelle, B., Alami, M., and Hopkins, C. R. (1993) ATP-dependent translocation of ricin across the membrane of purified endosomes. *J. Biol. Chem.* **268,** 23661–23669.
26. Sandvig, K., Garred, Prydz, K., Kozlov, J. V., Hansen, S. H., and van Deurs, B. (1992) Retrograde transport of endocytosed shiga toxin to the endoplasmic reticulum. *Nature* **358,** 510–512.
27. Stenmark, H., Moskaug, J., Madshus, I. H., Sandvig, K., and Olsnes, S. (1991) Peptides fused to the amino-terminal end of diphtheria toxin are translocated to the cytosol. *J. Cell Biol.* **113,** 1025–1032.
28. Madshus, I. H., Olsnes, S., and Stenmark, H. (1992) Membrane translocation of diphtheria toxin carrying passenger protein domains. *Infect. Immun.* **60,** 3296–3302.
29. Prior, T. I., FitzGerald, D. J., and Pastan, I. (1992) Translocation mediated by domain II of *Pseudomonas* exotoxin A: Transport of barnase into the cytosol. *Biochemistry* **31,** 3555–3559.

INFECTION AND IMMUNITY, June 1998, p. 2420–2425
0019-9567/98/$04.00+0

Vol. 66, No. 6

Biophysical Characterization of the Stability of the 150-Kilodalton Botulinum Toxin, the Nontoxic Component, and the 900-Kilodalton Botulinum Toxin Complex Species

FLORA CHEN,[1] GEOFFREY M. KUZIEMKO,[2,3] AND RAYMOND C. STEVENS[2,3]*

Graduate Group in Biophysics[1] and Department of Chemistry,[2] University of California, and Material Sciences Division, Lawrence Berkeley National Laboratory,[3] Berkeley, California 94720

Received 2 February 1998/Returned for modification 16 February 1998/Accepted 4 March 1998

Botulinum neurotoxin serotype A is initially released from the bacterium *Clostridium botulinum* as a stable 900-kDa complex. The serotype A 900-kDa complex is one of the forms of the toxin being used as a therapeutic agent for the treatment of various neuromuscular disorders. Previous experiments have demonstrated that the 900-kDa complex form of the toxin protects the toxin from the harsh conditions of the gastrointestinal tract. To provide molecular level details of the stability and equilibrium of the 900-kDa complex, the nontoxic component, and the toxic (botulinum neurotoxin) component, the three species have been investigated with a series of biophysical techniques at the molecular level (dynamic light scattering, proteolysis, circular dichroism, pH incubations, and agglutination assays). These experiments were conducted under harsh conditions which mimic those found along the gastrointestinal tract. Separately, exposure to denaturing and proteolytic conditions degrades both the botulinum neurotoxin and the nontoxic component. In the 900-kDa complex, the botulinum neurotoxin is protected during exposure to the gastrointestinal environment and the nontoxic component is slightly modified. Surprisingly, the toxin protects the ability of the nontoxic component to agglutinate erythrocytes. Contrary to previous reports, the purified 900-kDa complex did not have agglutination ability until after exposure to the proteolytic conditions. These experiments provide new evidence and detail for the theory that the nontoxic component and the toxic component protect one another during exposure to harsh conditions, and a molecular model is presented for the passage of the toxin through the gastrointestinal tract.

Botulinum neurotoxin is secreted by the anaerobic bacterium *Clostridium botulinum* as one of seven serotypes, classified A through G (35). The different serotypes proteolytically cleave specific proteins involved in synaptic vesicle docking that are necessary for cellular communication at the neuromuscular junction (12). Serotype A can be purified as a 900-kDa complex (BoNTA-HA) consisting of a toxic component (BoNTA [botulinum neurotoxin serotype A]) and a nontoxic component (HA [hemagglutinin]) (5–7). Studies involving relative oral toxicities (26, 33), intestinal absorption (34), and comparison to tetanus toxin (31) have indicated that the complex, not the botulinum toxin alone, is responsible for the extremely low amount of botulinum neurotoxin required in botulism poisoning.

The most common mechanism of botulism poisoning is through oral ingestion of the complex, which is found in food contaminated with *C. botulinum*. Ingested spores of the bacteria may also colonize and produce toxin in the intestinal tracts of infants, resulting in infant botulism (36). Previous experiments have demonstrated that the 900-kDa complex protects the toxin during its exposure to harsh conditions. Ohishi and coworkers (26) have demonstrated that the 900-kDa complex has a 360-fold-higher oral toxicity in rats than the 150-kDa botulinum neurotoxin component alone. Most proteins are broken down into peptides and amino acids in the stomach and small intestine during the process of digestion (2). However, the 900-kDa complex enters the stomach and withstands the acidic (pH 2) gastric juice containing the protease pepsin. The complex then enters the small intestine, where it encounters several more proteases (trypsin, chymotrypsin, and carboxypeptidase) that function at pH 7 to 8. Despite these harsh denaturing and proteolytic conditions, active botulinum neurotoxin (13, 22) and nontoxic component (34) can be detected in the lymph and circulatory systems.

The botulinum neurotoxin is comprised of a C-terminal 100-kDa heavy chain and an N-terminal 50-kDa light chain linked by a disulfide bond. From C terminus to N terminus, the protein can be further divided into three 50-kDa functional domains (3, 4, 18)—binding, translocation, and catalytic—which allow the protein to bind a cell surface receptor, pass across the membrane (29), and cleave a protein involved in synaptic vesicle docking (12), respectively. The nontoxic component is composed of several protein subunits and can be separated into nontoxic agglutinating proteins (14, 17, 32) and nontoxic nonagglutinating proteins (11, 14). Several subcomponents of the nontoxic component have been characterized. One of the nontoxic nonagglutinating proteins is a single peptide of 120 kDa and has been sequenced (11). The molecular weights of the subcomponents of the nontoxic agglutinating portion have also been determined (14, 32). However, the physiological role and molecular organization of the nontoxic component are not well understood. In addition, many of the studies of toxin exposure to harsh conditions have been conducted with crude cell supernatant. To understand the stability and equilibrium of the botulinum neurotoxin complex and its separated components at the molecular level, we conducted a series of biophysical experiments using pure preparations of botulinum neurotoxin, botulinum neurotoxin complex, and the nontoxic agglutinating portion.

* Corresponding author. Mailing address: Department of Chemistry, University of California, Berkeley, CA 94720. Phone: (510) 643-8285. Fax: (510) 643-9290. E-mail: stevens@adrenaline.berkeley.edu.

MATERIALS AND METHODS

Purification of proteins. For all experiments, proteins were obtained as ammonium sulfate precipitates and were purified by ion-exchange chromatography as previously described (5).

Light scattering. The hydrodynamic radius, estimated molecular weight, and polydispersity of protein samples were determined by using a Dynapro-801 dynamic light scattering instrument (Protein Solutions, Charlottesville, Va.). Samples of botulinum neurotoxin (0.18 mg/ml), nontoxic component (0.20 mg/ml), and 900-kDa complex (0.22 mg/ml) were incubated for 30 min at 4°C over a range of pHs in 10 mM buffers containing 100 mM sodium chloride. Buffers were citric acid (pH 1 to 5) (citric acid has three pK_as at 25°C; $pK_1 = 3.128$, $pK_2 = 4.761$, and $pK_3 = 6.396$), bis-Tris [bis(2-hydroxyethyl)imino-tris(hydroxymethyl) methane; pH 6], HEPES (pH 7 to 8), and CHES [2-(*N*-cyclohexylamino)ethanesulfonic acid] (pH 9 to 10). Individual samples were recorded for at least 5 min, during which 12 to 13 data points were taken and then analyzed by using a monomodal fitting program. The polydispersity served as the deviation of the size distribution based on the mean hydrodynamic radius of each sample. Experiments were performed in triplicate.

Test for pH effects. Protein samples were incubated at the appropriate pH for 48 h at 4°C. Protein samples were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 12% polyacrylamide gel, and the bands were visualized with silver stain (10).

Proteolysis. Proteins were incubated at various pHs and with several protease mixtures. Each protein sample contained 2 ml of protein (0.42 mg/ml) in dialysis tubing with a 50-kDa pore size. Dialysis tubing containing the protein sample was placed into the appropriate protease mixture. The mixture was incubated with gentle stirring for 2 h at 37°C. Pepsin incubations were performed with 0.4 mg of pepsin per ml in 10 mM hydrochloric acid (pH 2). Intestinal cocktail consisted of trypsin (0.02 mg/ml), chymotrypsin (0.03 mg/ml), and carboxypeptidase A (0.02 mg/ml) in 50 mM sodium bicarbonate–100 mM sodium chloride buffer (pH 8). Protease concentrations used were chosen based on reported values (15, 16), though the physiological protease concentration depends on the feeding state of the organism. After incubation, proteases were quenched with 0.5 ml of pepstatin (1 mg/ml) or phenylmethylsulfonyl fluoride (17.4 mg/ml) and 1 ml of 0.5 M EDTA. Proteins were also exposed to incubation of pepsin followed by intestinal cocktail or incubation at pH 2 followed by intestinal cocktail. Gels were 12% polyacrylamide, and protein bands were visualized with silver stain (10).

Circular dichroism. Circular dichroic spectra were gathered on a model J-600 spectropolarimeter (Jasco, Easton, Md.) at 37 and 25°C. The bandwidth used was 1 nm, and the step resolution was 2 nm. Four scans of each sample were made, using a time constant of 1 s and a scan rate of 50 nm/min. The cell volume was approximately 1 ml, with a path length of 0.1 cm. The cell was jacketed for temperature adjustment and controlled by using a water bath. Proteins at a concentration of 0.025 mg/ml were dialyzed extensively in their buffers before being examined on the spectropolarimeter. The pH 6 and pH 8 buffers were 100 mM sodium phosphate containing 100 mM sodium chloride. The pH 2 buffer was 60 mM hydrochloric acid-potassium chloride buffer containing 100 mM NaCl. No appreciable difference in signal was observed at 37 and 25°C; thus, only 37°C data are shown.

ELISA. Botulinum neurotoxin (0.01 mg/ml) was incubated in pepsin (0.4 mg/ml) and intestinal cocktail (see above) for 2 h at 37°C. The concentration of substrate protein used was the minimum amount of protein necessary to give strong enzyme-linked immunosorbent assay (ELISA) signal. Using single-chain variable fragments of mouse or of human monoclonal antibodies, ELISAs were performed on proteolyzed samples as described by Chen et al. (5).

Surface plasmon resonance. Binding of botulinum neurotoxin and 900-kDa complex to lipid was monitored by using surface plasmon resonance on a BIAcore 2000 (Biosensor, Piscataway, N.J.). Botulinum neurotoxin (84 to 1,300 nM) or 900-kDa complex (52 to 832.5 nM) was injected over a lipid monolayer containing 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC). The lipid monolayer was formed by incubating a solution of POPC liposomes (2 mg/ml) on a BIAcore sensorchip (type HPA) overnight in a moist environment. Liposomes were made by sonicating a solution of POPC in a buffer of 10 mM Tris (pH 7.0) containing 100 mM sodium chloride and 2 mM sodium azide. The experiments were performed at a flow rate of 40 ml/min for 20 min in running buffer of 10 mM Tris (pH 7.0) containing 100 mM sodium chloride and 2 mM sodium azide.

Sugar binding. Agglutination assays were performed as previously described (20). Erythrocytes (RBC) from rabbits were washed twice with saline (68 mM sodium citrate, 146 mM sodium chloride [pH 7.4]), spun at 4,000 × *g*, and diluted to 0.5% for each assay. Protein solution (0.5 ml) at twice the desired concentration was added to a disposable semimicrocuvette. Then 0.5 ml of the 0.5% RBC was added to the protein, resulting in the desired protein concentration and a 0.25% RBC solution. The mixture was incubated at 37°C for 2 h, and the transmittance was measured at 405 nm in a Shimadzu UV-160 spectrophotometer. The saline control contained 0.25% RBC incubated in saline for 2 h at 37°C. Each sample was performed in triplicate. The 900-kDa complex and nontoxic component were also proteolyzed and then used in an agglutination assay. The saline control was subtracted from the values for the other samples. The agglutination ability of the unproteolyzed HA was scaled to a value of 100. The agglutination abilities of the other samples represent their percent transmittances relative to the percent transmittance by unproteolyzed HA. Previous reports (20) show that the 900-kDa BoNTA-HA complex is capable of agglutination at lower temperatures. When we performed the agglutination assay at a lower temperature (12°C), we observed that the complex showed an increased ability to agglutinate RBC, consistent with the literature but not physiologically relevant.

RESULTS AND DISCUSSION

Light scattering. Dynamic light scattering was used to determine the aggregation state and stability of the three species under a variety of conditions simulating the environments in the gastrointestinal tract. This technique measures the diffusion coefficient of particles in solution, allowing a radius and estimated molecular weight to be calculated. A molecular weight calculated to be higher than expected indicates a deviation from a spherical shape. This deviation may be due to nonspherical species, denatured protein, or the presence of large aggregates of many proteins. The distribution of particle sizes is indicated by the polydispersity value. A polydispersity of zero signifies that there is only one particle size in solution. A large polydispersity value indicates the presence of different-size particles in solution. Examination of the radii of the three molecular species between pH 1 and 10 yielded a number of intriguing results. From the hydrodynamic radii the corresponding molecular weights were calculated, assuming an approximate, spherical shape for the protein.

Between pH 10 and 4, the estimated molecular mass of BoNTA was determined to be approximately 150 kDa (Fig. 1A and D), corresponding to the expected molecular mass of a single BoNTA molecule (6). Upon incubation at pH 3 to 1, botulinum neurotoxin showed steadily higher values for molecular mass, from 150 to over 1,000 kDa, during the first 10 min in the corresponding buffer. The polydispersity also increased dramatically as the pH changed from 3 to 1. These higher values for polydispersity correlated with aggregation and acid-catalyzed degradation of the neurotoxin molecule. This result is consistent with botulinum neurotoxin retaining little toxicity at pH 3 to 1 (33). Examination of the nontoxic component at various pHs revealed behavior similar to that of the botulinum neurotoxin (Fig. 1B and D). The nontoxic component did not aggregate or disassemble between pH 10 and 5 and had an apparent estimated molecular mass of approximately 962 kDa. The molecular mass calculated by light scattering differs from that calculated by subtracting the molecular mass of the BoNTA from the molecular mass of the BoNTA-HA complex because the nontoxic component most likely is not spherical. Since the light scattering device calculates a molecular mass by a formula that assumes the protein is a sphere, deviations in sphericality will result in deviations in molecular mass. The nontoxic component aggregated, disassembled, or denatured when the pH dropped from pH 5 to 4, corresponding to a sharp increase in the polydispersity between pH 5 (0%) and 4 (polydispersity/radius = 18.1%). Below pH 4, the nontoxic component increased in size to over 2,000 kDa. In contrast to the botulinum neurotoxin and the nontoxic component, the entire BoNTA-HA complex was most stable and monodisperse between pH 1 and 4 (Fig. 1C and D), correlating with the observation that the BoNTA-HA complex retains over 60% of its toxicity at low pH (33). Between pH 5 and 7, the complex reached its maximum size. Near neutral pH, the complex reached its maximum polydispersity. A similar high polydispersity was observed upon ultracentrifugation at pH 7.5 (37). Above pH 8, the complex appeared as a more polydisperse species, suggesting the presence of more than one species in solution probably due to disassociation of the complex into the toxic and nontoxic components. This idea of dissociation at basic pH is supported by the procedure for

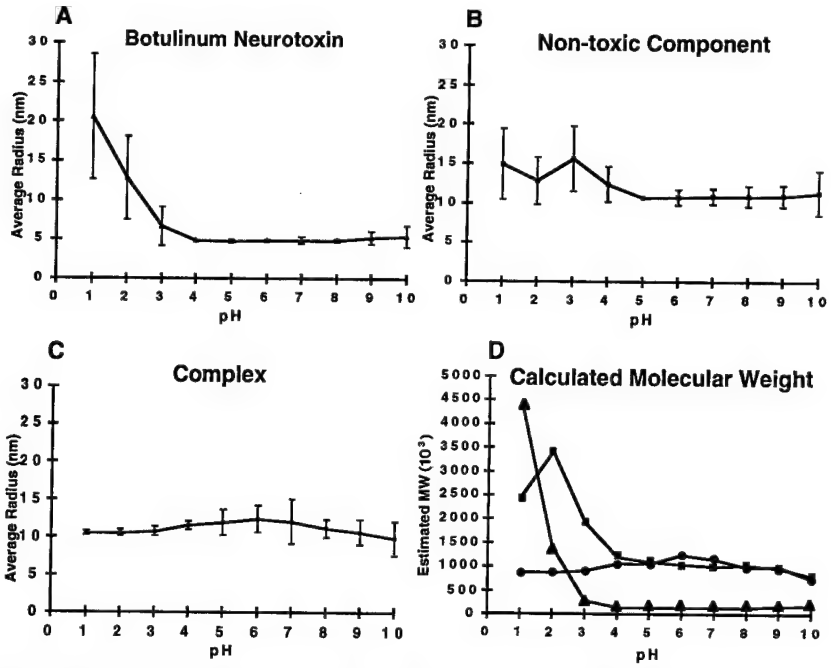


FIG. 1. Dynamic light scattering results, expressed as pH versus apparent radius and molecular weight of botulinum neurotoxin (A), nontoxic component (B), and 900-kDa complex (C) at various pHs. (D) Estimated molecular weights (MW) of botulinum neurotoxin (triangles), nontoxic component (squares), and 900-kDa complex (circles). The bar around each data point represents the average polydispersity value for three separate experiments. The polydispersity is the deviation of size distributions based on the mean radius measurement. Specimens containing aggregates or unfolded proteins have large polydispersity values.

purifying botulinum neurotoxin from the complex, since high pH is necessary for separation of toxic and nontoxic components (6, 7).

pH effects. The effect of pH on the neurotoxin's structural stability was examined by SDS-PAGE. A polyacrylamide gel of botulinum neurotoxin incubated at pH 2, 6, and 8 for 48 h indicated that the neurotoxin is susceptible to breakdown at pH 2 and 8 (Fig. 2). Table 1 shows that nonincubated neurotoxin runs as a 100-kDa heavy chain and a 50-kDa light chain. After 48 h of incubation at pH 6, we observed primarily the 100-kDa heavy chain and 50-kDa light chain of botulinum neurotoxin. A band at 150 kDa corresponded to a small quantity of unnicked neurotoxin. Although the 100- and 50-kDa bands predominated at pH 2 and 8, we observed a ladder of many lower-molecular-weight fragments. This fragmentation could be a result of acid-catalyzed hydrolysis, base-catalyzed hydrolysis, or proteolysis.

Proteolysis. Besides pH extremes, the other stresses that the 900-kDa complex must endure come from proteases. The effects of proteases on botulinum neurotoxin, nontoxic component, and 900-kDa complex were observed by performing protease incubations and assaying the results by gel electrophoresis (Table 1). Incubations consisted of pepsin at pH 2, intestinal cocktail (trypsin, chymotrypsin, carboxypeptidase) at pH 8, and pepsin at pH 2 followed by intestinal cocktail at pH 8. When incubated individually, both the neurotoxin and the nontoxic component were susceptible to all of the protease incubations (Table 1). In contrast, the 900-kDa complex demonstrated an amazing resistance to proteolysis. After incubation in pepsin at pH 2, the 900-kDa complex lost none of its bands. When the 900-kDa complex was incubated in intestinal cocktail, the 100- and 50-kDa bands of botulinum neurotoxin were proteolyzed along with the 120- and 106-kDa bands of the nontoxic component. However, incubating the 900-kDa complex in pepsin at pH 2 before exposing the complex to intestinal cocktail inhibited the proteolysis of the bands corresponding to botulinum neurotoxin. The bands which have been previously shown to be responsible for agglutination (14, 19, 21, 35, and 52 kDa) (14, 32) remained intact according to our gel whereas the band previously shown to be nontoxic-nonagglutinating (11) was destroyed.

Thus, under conditions that simulate the path of the 900-kDa complex through the gastrointestinal tract, the nontoxic component protected botulinum neurotoxin from proteolysis. This protection of the neurotoxin could be due to exposure to pepsin or to acidic pH. To determine if the protection of botulinum neurotoxin was a result of low pH, we performed an incubation without pepsin at pH 2, followed by an intestinal cocktail incubation. With these conditions, we observed that low pH was sufficient for protecting the complexed botulinum neurotoxin from proteolysis (data not shown). This protection of the botulinum neurotoxin in the 900-kDa complex was also seen upon incubations in rat gastric juice (33).

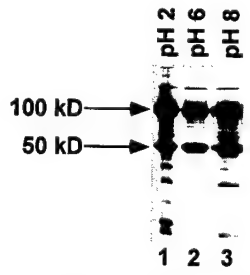


FIG. 2. SDS-PAGE of botulinum neurotoxin incubated at pH 2, 6, and 8. Arrows identify the heavy chain (100 kDa) and light chain (50 kDa) of botulinum neurotoxin. The band appearing at approximately 150 kDa is unnicked neurotoxin.

FIG. 3. Circular dichroism (CD) spectra of the 900-kDa complex at pH 8 (top), pH 6 (middle), and pH 2 (bottom) (A), the nontoxic component at pH 2 (top), pH 6 (middle), and pH 8 (bottom) (B), and botulinum neurotoxin at pH 8 (top), pH 6 (middle), and pH 2 (bottom) (C).

Circular dichroism. Circular dichroic spectra of the botulinum neurotoxin, the nontoxic component, and the complex were taken at 37 and 25°C to determine the effects of pH on the structure and stability of the proteins. Temperature appeared not to have an effect on the spectra. At pH 2, 6, and 8, the 900-kDa complex shows little change in secondary structure (Fig. 3A). Thus, the increased protease resistance of the 900-kDa complex, upon exposure to pH 2, did not result from a visible conformational change in the secondary structure. When the nontoxic component was studied at pH 6 and 8, it likewise showed minor change in secondary structure (Fig. 3B). Upon exposure to pH 2, the nontoxic component showed a slight decrease in helicity. This loss of secondary structure is not large enough to indicate that the nontoxic component significantly unfolds. Since individual nontoxic components do not unfold, an increase in hydrodynamic radius, which corresponds to an increase in the size of the particle being observed, likely results from the aggregation of multiple nontoxic components. This aggregation explains the size increase of the nontoxic component observed by light scattering (Fig. 1B).

The greatest change in secondary structure occurred with the botulinum neurotoxin. Circular dichroism spectroscopy on botulinum toxin and its separated components at pH 8.1 have shown that approximately 70% of the amino acid residues exist in an ordered structure (30). Previous reports of circular dichroic data from pH 9 to 6 revealed a slight increase in alpha helicity (8). Our results for botulinum neurotoxin show an increase in helicity on moving from pH 8 to 6 (Fig. 3C). However, a larger increase in helicity occurred on moving from pH 6 to 2. This increase in helicity may be due to a conformational change in the translocation domain of the neurotoxin, which is known to form pores upon exposure to acidic pH (4). If the neurotoxin increases its helicity at pH 2, producing more secondary structure, then botulinum neurotoxin does not unfold in an acidic environment. Consequently, the huge increase in

TABLE 1. Proteolysis and agglutination ability of BoNTA, nontoxic component (HA), and 900-kDa BoNTA-HA complex

Protein	Unproteolyzed		Pepsin (pH 2)		Intestinal cocktail (pH 8)		Sequential pepsin (pH 2)/ intestinal cocktail (pH 8)	
	No. of bands on 12% gel[a]	Agglutination[b]	No. of bands on 12% gel	Agglutination	No. of bands on 12% gel	Agglutination	No. of bands on 12% gel	Agglutination
BoNTA	2 (50, 100)	4.4	2 (30, 25)	—	4 (40, 45, 50, 100)	—	0	—
HA	7 (14, 19, 21, 35, 52, 106, 120)	100	7 (7, 14, 19, 21, 35, 52, 106, 120)	87.8	5 (14, 19, 21, 35, 52)	—	5 (14, 19, 21, 35, 52)	10.4
BoNTA-HA complex	9 (14, 19, 21, 35, 50, 52, 100, 106, 120)	1	9 (14, 19, 21, 35, 50, 52, 100, 106, 120)	9.4	5 (14, 19, 21, 35, 52)	—	7 (14, 19, 21, 35, 50, 52, 100)	37.9

[a] Numbers in parentheses indicate sizes (in kilodaltons) of bands appearing on 12% acrylamide gels under denaturing and reducing conditions.

[b] Percent transmittance by the protein relative to the percent transmittance by unproteolyzed HA, which was scaled to 100. Saline control contained 0.25% RBC incubated in saline for 2 h at 37°C. The saline control value was subtracted from the sample values before scaling to the unproteolyzed HA value. —, no activity.

hydrodynamic radius indicated by light scattering (Fig. 1A) is most likely caused by a conformational change in the molecular structure, followed by oligomerization or aggregation of the botulinum neurotoxin.

ELISA and surface plasmon resonance. The nontoxic component appears to prevent the vulnerable regions of the botulinum neurotoxin from protease attack, acid hydrolysis, and aggregation at low pH. Epitope mapping suggests that in the 900-kDa complex, the nontoxic component covers a large portion of the binding domain of botulinum neurotoxin (5). Previous experiments have shown that the binding domain is highly susceptible to trypsin cleavage (28). Therefore, the nontoxic component may play an important role in protecting the binding domain in the gastrointestinal tract. Furthermore, ELISA performed on botulinum neurotoxin after pepsin and intestinal proteolysis showed that antibodies no longer bind to it, suggesting that cleavage, denaturation, or both have occurred. In addition to being highly susceptible to proteolysis, botulinum neurotoxin appears to stick strongly and irreversibly to lipid monolayers such as POPC, a lipid commonly found in membranes of cells lining the gastrointestinal tract (9). Using surface plasmon resonance, we found that botulinum neurotoxin injected over POPC monolayers adsorbed strongly whereas the 900-kDa complex did not. Similarly, unless the toxic component was in its complexed form, it would probably adhere to lipid membranes along the gastrointestinal tract. Additional evidence for the association of botulinum neurotoxin with phospholipids has been seen by several researchers (17, 24, 25, 27). Thus, without the nontoxic component for protection, botulinum neurotoxin might never leave the gastrointestinal tract but instead be digested like most other proteins.

Sugar binding. The only known ability of the nontoxic component is its capacity to bind certain sugars (1). Upon binding these sugars, the nontoxic component is able to agglutinate, or clump together, cells (19). Whether this property aids the neurotoxin in reaching its target is under investigation. The clumping of RBC was monitored with agglutination assays (Table 1). Previous work has shown that the 14-, 19-, 21-, 35-, and 52-kDa bands are primarily responsible for agglutination (14, 32). Unproteolyzed nontoxic component containing these five bands was able to agglutinate RBC at 37°C (Table 1). After proteolysis by pepsin (pH 2), the nontoxic component experienced a slight decrease in agglutination ability (Table 1). Upon exposure to sequential protease incubations, the nontoxic component showed the largest decrease in the ability to agglutinate RBC (Table 1). This decrease appears to correlate with the loss of the 120- and 106-kDa bands (Table 1). The 106-kDa band was shown to be a proteolyzed portion of the 120-kDa band (14). Therefore, the 120-kDa band appears to be necessary for the nontoxic component to maintain optimal agglutination ability by protecting the other proteins but is not directly involved in agglutination. When unproteolyzed or exposed to pepsin, the complex showed a low ability to agglutinate (Table 1). Upon exposure to a sequential pepsin and intestinal cocktail incubation, the complex increased its ability to agglutinate RBC (Table 1). The sequentially proteolyzed complex contains the 14-, 19-, 21-, 35-, and 52-kDa bands plus the 100- and 50-kDa bands corresponding to botulinum neurotoxin (Table 1). Thus, the presence of botulinum neurotoxin appears to protect the ability of the complex to agglutinate after a sequential protease incubation.

These experiments indicate that in order for the nontoxic component to maintain optimal agglutination activity, it must be part of the 900-kDa complex while traveling through the gastrointestinal tract. After exposure to the conditions simulating the gastrointestinal tract, the 900-kDa complex appears to be modified so that the agglutination ability is activated. We propose that the 900-kDa complex may then interact with sugars on RBC and release the 150-kDa neurotoxin into the circulatory system. Although further experiments are necessary, experiments in which the 900-kDa complex was incubated with RBC that were subsequently pelleted by centrifugation and washed in saline solution (0.85% NaCl) showed that the neurotoxin was released and remained active in the supernatant (21). In addition, the neurotoxin has been shown to separate from the nontoxic component when the latter is bound to a sugar affinity column (23) and eluted with buffers at basic pH and ionic strength.

The 900-kDa complex form of botulinum neurotoxin is necessary for the delivery of botulinum neurotoxin in its most potent form. Using purified material and examining the results of the different toxin species at the molecular level, we have provided further evidence that the neurotoxin must exist in the 900-kDa complex to maintain its activity in conditions mimicking the environment of the gastrointestinal tract. In addition, a synergistic partnership between botulinum neurotoxin and the nontoxic component in which the nontoxic portion preserves the toxic ability of the toxic portion and the toxic portion protects the agglutinating ability of the nontoxic component seems to exist. In the 900-kDa complex, the neurotoxin and nontoxic component protect each other from pH extremes and gastrointestinal proteases. In contrast, when separated and exposed to simulated digestive conditions, each component is degraded. Although the molecular mechanism of the neurotoxin's journey to the neuromuscular junction is still unclear, these biophysical studies provide further evidence and detail as to the importance of the 900-kDa complex in the potent oral toxicity of botulinum neurotoxin.

ACKNOWLEDGMENTS

F. Chen and G. M. Kuziemko contributed equally to the work.

We thank B. R. DasGupta and Bill Tepp for generous supplies of botulinum neurotoxin, 900-kDa complex, and nontoxic component. We also thank the lab of Ignacio Tinoco, Jr., especially Barbara Dengler, for use of their circular dichroism spectropolarimeter.

This work was supported in part by the U.S. Army Medical Research and Development Command (DAMD17-93-C-3118), an NSF Young Investigator Award (R.C.S.), U.S. Department of Energy contract DE-AC03-76SF0098, and a National Institutes of Health predoctoral traineeship from the Neurobiology Division, University of California, Berkeley.

REFERENCES

1. **Balding, P., E. R. Gold, D. A. Boroff, and T. A. Roberts.** 1973. Observations on receptor specific proteins. II. Haemagglutination and haemagglutination-inhibition reactions of Clostridial botulinum types A, C, D and E haemagglutinins. Immunology **25:**773–782.
2. **Beck, W. S., K. F. Liem, and A. R. Simpson.** 1991. Life, an introduction to biology, 3rd ed., p. 666–676. Harper Collins Publishers, Inc., New York, N.Y.
3. **Blasl, J., E. R. Chapman, E. Link, T. Binz, S. Yamasaki, P. De Camill, T. C. Südhof, H. Niemann, and R. Jahn.** 1993. Botulinum neurotoxin A selectively cleaves the synaptic protein SNAP-25. Nature **365:**160–163.
4. **Blaustein, R. O., W. J. Germann, A. Finkelstein, and B. R. DasGupta.** 1987. The N-terminal half of the heavy chain of botulinum type A neurotoxin forms channels in planar phospholipid bilayers. FEBS Lett. **226:**115–120.
5. **Chen, F., G. M. Kuziemko, P. Amersdorfer, C. Wong, J. D. Marks, and R. C. Stevens.** 1997. Antibody mapping to domains of botulinum neurotoxin serotype A in the complexed and uncomplexed forms. Infect. Immun. **65:** 1626–1630.
6. **DasGupta, B. R., and D. A. Boroff.** 1968. Separation of toxin and hemagglutinin from crystalline toxin of *Clostridium botulinum* type A by anion exchange chromatography and determination of their dimensions by gel filtration. J. Biol. Chem. **243:**1065–1072.
7. **DasGupta, B. R., D. A. Boroff, and E. Rothstein.** 1966. Chromatographic fractionation of the crystalline toxin of *Clostridium botulinum* type A. Biochem. Biophys. Res. Commun. **22:**750–756.
8. **Datta, A., and B. R. DasGupta.** 1988. Circular dichroic and fluorescence

spectroscopic study of the conformation of botulinum neurotoxin types A and E. Mol. Cell. Biochem. **79:**153–159.
9. **Dudeja, P. K., J. M. Harig, K. Ramaswamay, and T. A. Brasitus.** 1989. Protein-lipid interaction in human small intestinal brush-border membranes. Am. J. Physiol. **257:**G809–G817.
10. **Fling, S. P., and D. S. Gregerson.** 1988. Peptide and protein molecular weight determination by electrophoresis using a high-molarity Tris buffer system without urea. Anal. Biochem. **155:**83–88.
11. **Fujita, R., Y. Fujinaga, K. Inoue, H. Nakajima, H. Kumon, and K. Oguma.** 1995. Molecular characterization of two forms of nontoxic-nonhemagglutinin components of Clostridium botulinum type A progenitor toxins. FEBS Lett. **376:**41–44.
12. **Hayashi, T., H. McMahon, S. Yamasaki, T. Binz, Y. Hata, T. C. Südhof, and H. Niemann.** 1994. Synaptic vesicle membrane fusion complex: action of clostridial neurotoxins on assembly. EMBO J. **2:**5051–5061.
13. **Heckly, R. J., G. J. Hildebrand, and C. Lamanna.** 1960. On the size of the toxic particle passing the intestinal barrier in botulism. J. Exp. Med. **111:** 745–759.
14. **Inoue, K., Y. Fujinaga, T. Watanabe, T. Ohyama, K. Takeshi, K. Moriishi, H. Nakajima, K. Inoue, and K. Oguma.** 1996. Molecular composition of *Clostridium botulinum* type A progenitor toxins. Infect. Immun. **64:**1589–1594.
15. **Johnson, L. R.** 1981. Physiology of the gastrointestinal tract, p. 625–627. Raven Press Books, Ltd., New York, N.Y.
16. **Johnson, L. R.** 1981. Physiology of the gastrointestinal tract, p. 784–787. Raven Press Books, Ltd., New York, N.Y.
17. **Kamata, Y., Y. Kimura, and S. Kozaki.** 1994. Involvement of phospholipids in the intoxication mechanism of botulinum neurotoxin. Biochim. Biophys. Acta **1199:**65–68.
18. **Kozaki, S., A. Miki, Y. Kamata, T. Nagai, J. Ogasawara, and G. Sakaguchi.** 1989. Immunological characterization of the papain-induced fragments of *Clostridium botulinum* type A neurotoxin and interaction of the fragments with brain synaptosomes. Infect. Immun. **57:**2634–2639.
19. **Lamanna, C.** 1948. Hemagglutination by botulinal toxin. Proc. Soc. Exp. Biol. Med. **69:**332–336.
20. **Lowenthal, J. P., and C. Lamanna.** 1951. Factors affecting the botulinal hemagglutination reaction, and the relationship between hemagglutinating activity and toxicity of toxin preparations. Am. J. Hyg. **54:**342–353.
21. **Lowenthal, J. P., and C. Lamanna.** 1953. Characterization of botulinal hemagglutination. Am. J. Hyg. **57:**46–59.
22. **May, A. J., and B. C. Whaler.** 1958. The absorption of *Clostridium botulinum* type A toxin from the alimentary canal. Br. J. Exp. Pathol. **39:**307–316.
23. **Moberg, L. J., and H. Sugiyama.** 1978. Affinity chromatography purification of type A botulinum neurotoxin from crystalline toxin complex. Appl. Environ. Microbiol. **35:**878–880.
24. **Montecucco, C., G. Schiavo, Z. Gao, E. Bauerlein, P. Boquet, and B. R. DasGupta.** 1988. Interaction of botulinum and tetanus toxins with the lipid bilayer surface. Biochem. J. **251:**379–383.
25. **Montecucco, C., G. Schiavo, and B. R. DasGupta.** 1989. Effect of pH on the interaction of botulinum neurotoxins A, B and E with liposomes. Biochem. J. **259:**47–53.
26. **Ohishi, I., S. Sugii, and G. Sakaguchi.** 1977. Oral toxicities of *Clostridium botulinum* in response to molecular size. Infect. Immun. **16:**107–109.
27. **Schmid, M. F., J. P. Robinson, and B. R. DasGupta.** 1993. Direct visualization of botulinum neurotoxin-induced channels in phospholipid vesicles. Nature **364:**827–830.
28. **Shone, C. C., P. Hambleton, and J. Melling.** 1985. Inactivation of Clostridium botulinum type A neurotoxin by trypsin and purification of two tryptic fragments. Eur. J. Biochem. **151:**75–82.
29. **Simpson, L. L.** 1980. Kinetic studies on the interaction between botulinum toxin type A and the cholinergic neuromuscular junction. J. Pharmacol. Exp. Ther. **212:**16–21.
30. **Singh, B. R., and B. R. DasGupta.** 1989. Structure of heavy and light chain subunits of type A botulinum neurotoxin analyzed by circular dichroism and fluorescence measurements. Mol. Cell. Biochem. **85:**67–73.
31. **Singh, B. R., B. Li, and D. Read.** 1995. Botulinum versus tetanus neurotoxins: why is botulinum neurotoxin but not tetanus neurotoxin a food poison? Toxicon **33:**1541–1547.
32. **Somers, E., and B. R. DasGupta.** 1991. *Clostridium botulinum* types A, B, C_1, and E produce proteins with or without hemagglutinating activity: do they share common amino acid sequences and genes? J. Protein Chem. **10:**415–425.
33. **Sugii, S., I. Ohishi, and G. Sakaguchi.** 1977. Correlation between oral toxicity and in vitro stability of *Clostridium botulinum* type A and B toxins of different molecular sizes. Infect. Immun. **16:**910–914.
34. **Sugii, S., I. Ohishi, and G. Sakaguchi.** 1977. Intestinal absorption of botulinum toxins of different molecular sizes in rats. Infect. Immun. **17:**491–496.
35. **Sugiyama, H.** 1980. *Clostridium botulinum* neurotoxin. Microbiol. Rev. **44:** 419–448.
36. **Tacket, C. O., and M. A. Rogawski.** 1989. Botulism, p. 354-356. *In* L. L. Simpson (ed.), Botulinum neurotoxin and tetanus toxin. Academic Press, Inc., San Diego, Calif.
37. **Wagman, J., and J. B. Bateman.** 1953. Botulinum type A toxin: properties of a toxic dissociation product. Arch. Biochem. Biophys. **45:**375–383.

Editor: J. T. Barbieri

Crystal structure of botulinum neurotoxin type A and implications for toxicity

D. Borden Lacy[1], William Tepp[2], Alona C. Cohen[1], Bibhuti R. DasGupta[2] and Raymond C. Stevens[1]

Botulinum neurotoxin type A (BoNT/A) is the potent disease agent in botulism, a potential biological weapon and an effective therapeutic drug for involuntary muscle disorders. The crystal structure of the entire 1,285 amino acid di-chain neurotoxin was determined at 3.3 Å resolution. The structure reveals that the translocation domain contains a central pair of α-helices 105 Å long and a ~50 residue loop or belt that wraps around the catalytic domain. This belt partially occludes a large channel leading to a buried, negative active site — a feature that calls for radically different inhibitor design strategies from those currently used. The fold of the translocation domain suggests a mechanism of pore formation different from other toxins. Lastly, the toxin appears as a hybrid of varied structural motifs and suggests a modular assembly of functional subunits to yield pathogenesis.

The clostridial neurotoxin family is composed of tetanus neurotoxin and seven serotypes of botulinum neurotoxin (BoNT/A–BoNT/G). Botulinum neurotoxin serotype A is synthesized in *Clostridium botulinum* as a ~150,000 M_r single chain protein (1,296 amino acids). The toxin is post-translationally proteolyzed to form a 1,285 amino acid di-chain molecule in which the two chains, ~50,000 M_r and ~100,000 M_r, remain linked by a disulfide bond[1]. The protein is composed of three ~50,000 M_r functional domains where the catalytic function is confined to the ~50,000 M_r chain (residues 1–437), the translocation activity with the N-terminal half of the ~100,000 M_r chain (residues 448–872), and receptor binding with its C-terminal half (residues 873–1,295)[1,2]. The toxicity of BoNT is a result of a multi-step mechanism[3]. The neurotoxin binds to the pre-synaptic nerve endings of cholinergic neurons and enters by receptor-mediated endocytosis. Acidity in the endosome is believed to induce pore formation, which allows translocation of the catalytic domain into the cytosol. The zinc-dependent catalytic domain of the seven BoNT serotypes specifically cleaves one of three different SNARE proteins essential for synaptic vesicle fusion: synaptobrevin, syntaxin, or SNAP-25. This cleavage results in inhibition of acetylcholine secretion, ultimately leading to paralysis[1].

What remains unknown is how the highly homologous BoNT serotypes and tetanus toxin derive specificity both for different receptors and their different SNARE protein targets. Furthermore, a number of questions surround the mechanism by which a soluble protein changes with acidic pH to one that can span a membrane and translocate a ~50,000 M_r protein into the cytoplasm . We describe here the crystal structure of BoNT/A and then use this structural information to address some of these mechanistic questions.

Overall structure of BoNT/A

The crystal structure of BoNT/A was determined by multiple isomorphous replacement (MIR) using five heavy-atom derivatives (Table 1). Phases were further improved with solvent flattening to yield readily interpretable electron density maps (Fig. 1). The model, which includes 99% of the amino acids, was refined to 3.3 Å with an R-value of 20.0% and an R_{free} value of 27.9%. The BoNT/A molecule is approximately 45 Å × 105 Å × 130 Å and shows a linear arrangement of the three functional domains with no contact between the catalytic and binding domains (Fig. 2). In general, the three functional domains are also structurally distinct. The exception is an unusual loop, or belt, from what has historically been considered part of the translocation domain, that wraps around the perimeter of the catalytic domain.

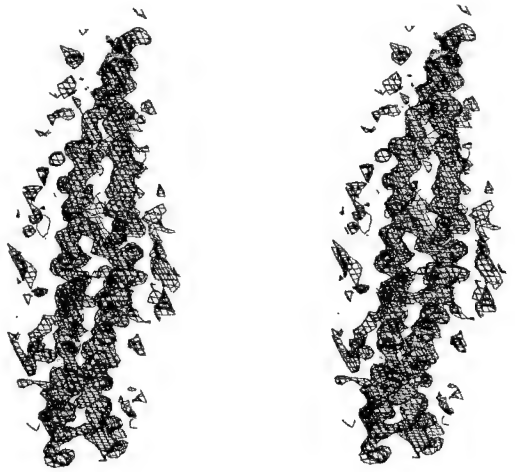


Fig. 1 Stereo diagram of the initial electron density. The map was calculated with MIR phases extended to 3.6 Å, improved by density modification, and contoured at 2.0σ. This region of the structure corresponds to the pair of α-helices in the translocation domain.

[1]Department of Chemistry and Earnest Orlando Lawrence Berkeley National Laboratory, University of California, Berkeley, California 94720, USA. [2]Department of Food Microbiology and Toxicology, University of Wisconsin, Madison, Wisconsin 53706, USA.

Correspondence should be addressed to R.C.S. *e-mail: stevens@adrenaline.berkeley.edu*

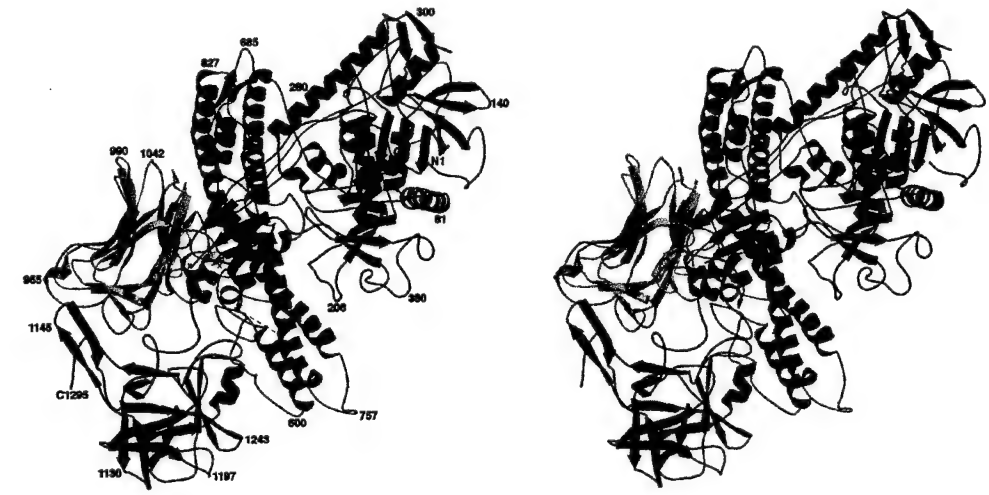


Fig. 2 Stereo diagram of the backbone trace for the BoNT/A model. The catalytic domain is colored in blue, the translocation domain in green, the N-terminal binding sub-domain in yellow and the C-terminal binding sub-domain in red. The catalytic zinc is depicted as a ball in gray. The 'X' dictates the proteolytic site separating the two chains. The dashed line is the only loop region with ill-defined electron density. The lower view shows the model rotated ~90° out of the plane of the page relative to the top view and should better depict the translocation domain belt wrapping around the catalytic domain. The overall structure is 45 Å × 105 Å × 130 Å. Figure generated using MOLSCRIPT[36].

Receptor binding domain

The binding domain, with overall dimensions of 32 Å × 37 Å × 76 Å, appears as two distinct sub-domains roughly equal in size. Both sub-domains are composed predominantly of β-strands and are connected by one prominent α-helix. The N-terminal sub-domain has two seven-stranded β-sheets sandwiched together in a jelly roll motif with dimensions 32 Å × 37 Å × 38 Å. The C-terminal sub-domain has similar dimensions and adopts a modified β-trefoil fold with a six-stranded β-barrel next to the N-terminal jelly roll motif and a β-hairpin triplet capping the base of the domain. The entire binding domain tilts away from the long helical axis of the translocation domain such that there is no contact between the C-terminal sub-domain and the translocation domain and all of the surface loops are accessible for binding. The binding domain shares structural homology with the recently solved structure of the tetanus toxin binding domain[4] with an r.m.s. deviation of 1.5 Å for 363 Cα positions (Fig. 3). The major differences appear in the loops of the C-terminal sub-domain, where the longer loop lengths in the tetanus toxin sub-domain account for the domain's slightly longer length in primary sequence.

The first step in the intoxication mechanism is a binding event between the binding domain and the pre-synaptic nerve ending. The interaction with BoNT/A is proposed to occur through both a polysialoganglioside (G_{D1b} or G_{T1b}) and a yet unidentified protein receptor[5,6]. The putative ganglioside binding site for tetanus neurotoxin, as assessed by photoaffinity labelling[7], is in a loop of the C-terminal sub-domain, and when overlayed with the BoNT/A holotoxin structure, is fully accessible. This could represent a general ganglioside binding site for all the clostridial neurotoxins. It is striking that both sub-domains have structural homology with proteins known to interact with sugars — as assessed using Dali[8], a three-dimensional search algorithm — and its ranked output according to Z-score. (The Z-score is the strength of structural similarity in

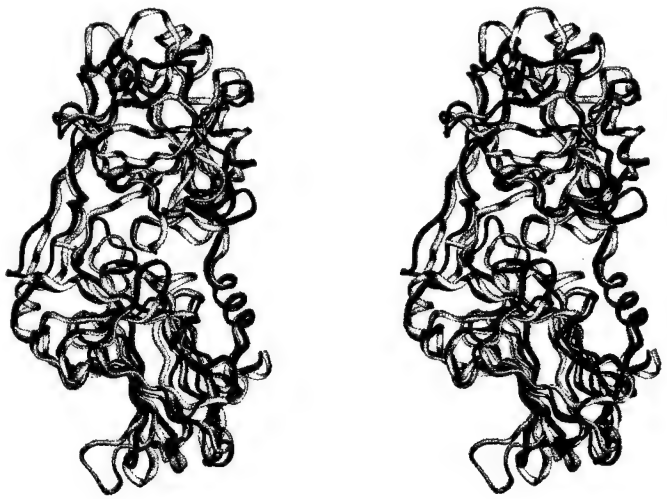


Fig. 3 Stereo superposition of the BoNT/A binding domain (residues 873–1,295) in red and the tetanus toxin binding domain (residues 874–1,314) in yellow. The root-mean-square (r.m.s.) deviation is 1.5 Å for 363 Cα atoms. The major structural differences are in the C-terminal sub-domain loops.

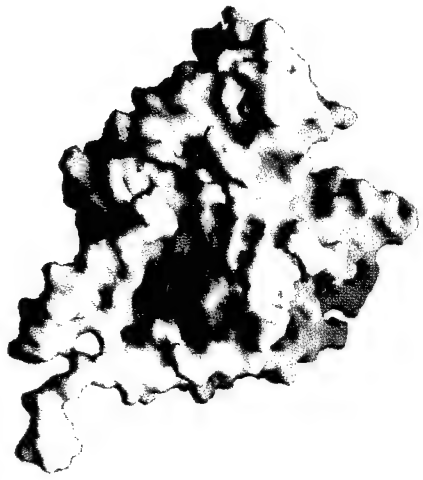


Fig. 4 Molecular surface of the catalytic domain colored by electrostatic potential (red = negative, blue = positive, white = uncharged). The active site is buried within the channel and appears largely negative. In the holotoxin, this pocket is partially blocked by the translocation domain belt. Figure generated using GRASP[37].

standard deviations over the expected where pairs with $Z < 2$ are structurally dissimilar.) The top twelve proteins after tetanus toxin ($Z = 26.3$) that are structurally similar to the N-terminal sub-domain are proteins known to interact with sugars (for example, serum amyloid P (1sac-A, $Z = 12.8$), β-glucanase (2ayh, $Z = 11.1$), sialidase (1kit, $Z = 8.9$), and lectin (1led, $Z = 8.8$)). Perhaps the most notable of these twelve are cryia (1ciy, $Z = 5.2$) and insecticidal ∂-endotoxin (1dlc, $Z = 4.8$) which act by binding glycoproteins and creating leakage channels[9]. These toxins have binding domains with motifs similar to the N-terminal sub-domain of BoNT/A but dramatically different pore-forming domains. A Dali search of the C-terminal sub-domain reveals more sugar binding proteins (basic fibroblast growth factor (1bfg, $Z = 9.1$), agglutinin (1jly-A, $Z = 9.9$)) and the toxin abrin (1abr-B, $Z = 8.8$). Both abrin and the related ricin bind their targets through a β-trefoil binding domain. The appearance of different subsets of the same structural motifs in different toxins could suggest a mechanism of evolution in which stable functional domains are assembled as modular units giving rise to toxicity.

Pore-formation and translocation domain

Following cell surface binding and receptor-mediated endocytosis of the neurotoxin, an acid-induced conformational change in the neurotoxin's translocation domain is believed to allow the translocation domain to penetrate the endosome and form a pore. The membrane interaction and pore formation is thought to facilitate the passage of the catalytic domain across the membrane into the cytosol. The details of how the translocation domain changes conformation at acidic pH to form a pore and how it can allow for the passage of a 50,000 M_r catalytic domain across the endosomal membrane are the least understood aspects of the intoxication mechanism. To date, most investigations of the translocation event in BoNT have assumed a similarity to other pH dependent α-helical pore-forming proteins: diphtheria toxin, colicin A, δ-endotoxin, pseudomonas exotoxin, and Bcl-x_L[9–13]. These proteins share a common structural motif in their pore forming domain[13,14], however this motif is not observed in the BoNT/A structure.

The translocation domain of BoNT/A wraps around the catalytic domain before forming its main body, a cylindrical shape with dimensions of 28 Å × 32 Å × 105 Å (Fig. 2). The most salient feature of the translocation domain is a pair of α-helices 105 Å long corresponding to residues 685–827. While unusual, long pairs of α-helices have been observed recently in the structures of colicin Ia[15] and the nucleotide exchange factor GrpE[16]. The helices, anti-parallel and amphipathic, twist around each other like a coiled coil but do not adhere to a strict heptad repeat. At both ends of this pair of helices, a shorter α-helix packs in parallel to the long helical axis. In addition, the domain has two strand-like sections which pack against the pair of α-helices in a parallel fashion. In an effort to identify the pore-forming segment of BoNT/A, the primary sequence was searched for predicted amphipathic helicity, and the candidate segment (residues 659–681) was shown to increase permeability of lipid bilayers[17]. The present structure, solved at neutral pH, shows that none of this putative transmembrane segment is helical and that, in fact, part of it appears to be in one of the two strand-like segments packing against the long α-helices. This could indicate an area that will undergo structural changes with pH. However, the

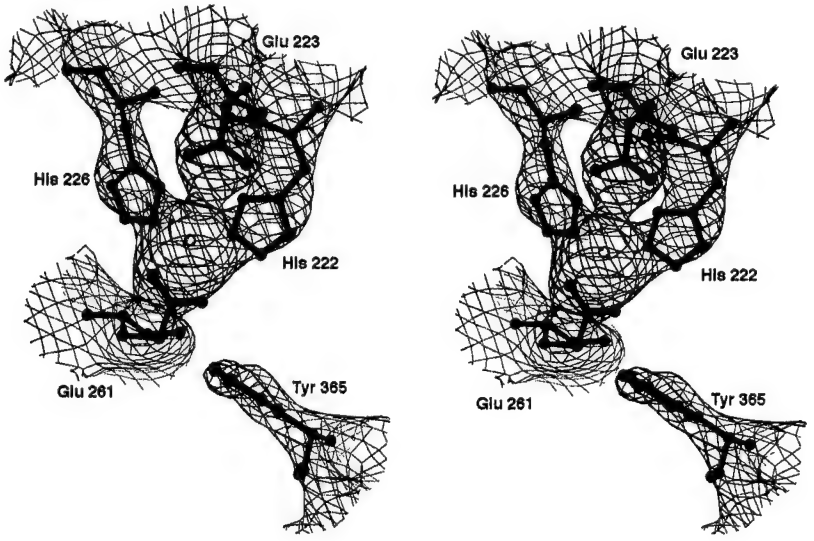


Fig. 5 Stereo diagram of the catalytic domain active site. A σ_A-weighted F_{obs} - F_{calc} omit map contoured at 2.6σ and centered around the zinc shows the amino acids (green carbon, red oxygen, and blue nitrogen bonds) thought to be critical for full catalytic activity. Figure generated using BOBSCRIPT.

Table 1 Structure determination statistics

Diffraction data statistics

Crystal	Resolution (Å)	Total obs.	Unique refl.	Completeness[1] (%)	R_{merge}[2] (%)
Native	20–3.3	172,169	40,110	98.2(94.3)	11.1(48.3)
CH_3HgCl	15–4.5	87,457	15,830	97.2(96.0)	10.7(27.7)
$Hg_2(Ac)_2$	15–4.2	86,260	19,650	97.8(94.3)	9.1(24.5)
SmAc	15–4.5	22,797	12,516	78.5(78.2)	7.6(17.4)
UAc	15–4.5	22,795	11,796	71.9(73.6)	7.8(22.5)
$KAuCl_4$	15–6.0	19,894	6,080	93.7(95.5)	9.4(18.2)

Phasing statistics

Derivative	Number of sites[3]	Isomorphous R- factor[4] (%)	Phasing power[5] Cent.	Phasing power[5] Acent.	R_{Cullis}[6] Cent.	R_{Cullis}[6] Acent.
CH_3HgCl	6	21.7	1.03	1.64	0.68	0.65
$Hg_2(Ac)_2$	6	21.0	0.96	1.36	0.76	0.76
SmAc	3	14.9	0.77	1.07	0.83	0.87
UAc	3	15.2	0.66	1.04	0.88	0.91
$KAuCl_4$	1	15.2	0.63	1.03	0.88	0.88
FOM before solvent flattening (15–3.6 Å)		0.48				
FOM after solvent flattening (15–3.6 Å)		0.74				

Refinement statistics

R-factor/R_{free}	20.0/27.9
R.m.s.d. bond lengths	0.008 Å
R.m.s.d. bond angles	1.5°

[1]Numbers in parentheses indicate statistics for highest resolution shells.
[2]$R_{merge} = \Sigma_h\Sigma_i | I_{hi} - \langle I_{hi}\rangle | / \Sigma_h\Sigma_i | I_{hi} |$, where h specifies unique indices, i indicates symmetry equivalent observations of h, and $\langle I_{hi}\rangle$ is the mean value.
[3] The mercury and gold derivatives shared overlapping sites. Similarly, the samarium and uranyl derivatives shared sites. However, the different occupancies yielded enough new phase information to improve the quality of the maps.
[4]The isomorphous R-factor = $\Sigma_h || F_{PH} | - | F_P || / \Sigma_h | F_P |$, where $| F_{PH} |$ and $| F_P |$ are the measured structure factor amplitudes of the derivative and native structures.
[5]Phasing power is the mean value of the heavy atom structure factor amplitude divided by the residual lack of closure error.
[6]R_{Cullis} is the mean residual lack of closure error divided by the isomorphous difference.

residues most likely to titrate over this pH range are the two histidines in the translocation domain, which are located some distance from this region of the structure: His 551 and His 560 are found in a loop between the translocation domain belt and the main body of the translocation domain. This junction may play a role in exposing a hydrophobic segment of the protein or releasing the catalytic domain from the translocation domain. Regardless of the pore-forming segment location, it is clear that the translocation domain of BoNT/A is structurally distinct from the other pore-forming toxins. In fact, the long pair of α-helices with their triple helix bundles at either end bear more resemblance to the coiled coil viral proteins: HIV-1 gp41/GCN4, influenza hemagglutinin, and the MoMuLV TM fragment[18–20]. These proteins do not translocate through pores but do have an acid-induced ability to undergo structural changes and penetrate membranes.

Catalytic domain

Either within the acidic endosome or upon exposure to the cytosol following translocation, the disulfide bond connecting the catalytic and translocation domains (Cys 429–Cys 453) is reduced, and the catalytic domain is released into the cytoplasm. The catalytic domain, 55 Å × 55 Å × 62 Å, is a mixture of both α-helix and β-strand secondary structure, in agreement with secondary structure predictions[21]. The active site of the catalytic domain is buried 20–24 Å deep in the protein, has a negative surface charge, and is accessible by a channel, ~12 Å × 15 Å × 35 Å, (Fig. 4). In the di-chain holotoxin, this channel is partially shielded from solvent by both the belt and the main body of the translocation domain.

The final mechanistic step in toxicity is the cleavage of a presynaptic protein by the BoNT zinc protease catalytic domain. The catalytic zinc atom represents the highest peak in the electron density maps and is visible in MIR maps contoured at 7.5σ. Amino acids with side chains closest to the zinc include His 222, Glu 223, His 226, Glu 261 and Tyr 365 (Fig. 5). While the presence of His 222, Glu 223 and His 226 was anticipated[1], this structure identifies Glu 261 as the fourth ligand, and may help resolve uncertainty in tetanus toxin where either Glu 269 or Glu 270 had been implicated[22]. (Glu 269 and Glu 270 of tetanus toxin align to Glu 260 and Glu 261 of BoNT/A). Further, while the role of a tyrosine was anticipated[23,24], the structure indicates that it is the conserved Tyr 365 and not the conserved Tyr 232 which is within proximity to the zinc. While exact bond lengths and water molecules cannot be confirmed at this resolution, the observed orientation of these residues support a model in which the His 222, His 226 and Glu 261 directly coordinate the zinc, and Glu 223 coordinates a water molecule as the fourth ligand. The Tyr 365 residue is ~5 Å from the zinc (OH–Zn) and is more likely to be involved in secondary bonding networks or interaction with substrate. A Dali search[8] shows that the proteins with the most structural similarity to the catalytic domain are thermolysin (1hyt, Z = 4.6) and leishmanolysin (1lml, Z = 2.4), two other zinc proteases with the same conserved HEXXH sequence. The structural similarities are limited, however, to the helix containing the HEXXH sequence and a four-stranded β-sheet buttressing the helix. Beyond this overlap, the BoNT/A catalytic domain has different secondary structure elements and connectivities.

The catalytic domain of BoNT/A is highly specific for the C-terminus of SNAP-25 and appears to require a minimum of

16 substrate amino acids for cleavage[25]. This large substrate size requirement is unusual for metalloproteases. The molecular recognition properties between the toxin and SNAP-25 and the dependence of recognition on a specific substrate secondary structure is presently being investigated[26,27]. The channel (~12 Å × 15 Å × 35 Å) leading to the active site (Fig. 4) appears capable of accommodating 16 or more residues with the Gln–Arg cleavage site of SNAP-25 able to insert into the zinc catalytic site. Additionally, the high pI of SNAP-25's C-terminal tail suggests that the negatively charged surface observed near the active site could be critical in the docking of substrate. Interestingly, the active site is most accessible in the absence of the translocation domain. This structure, in which the translocation domain shields the active site in the unreduced holotoxin, explains the fact that the catalytic activity in *in vitro* experiments is greatly enhanced by reduction of the disulfide. This observation of an occluded active site, along with the characterization of the amino acid environment, will almost certainly benefit ongoing inhibitor design.

Methods

Crystals were grown as described[28], but with a modified mother liquor of 150 mM magnesium acetate, 11% PEG 4000, and 100 mM tris-HCl pH 7. The crystals contain one monomer per asymmetric unit and are in space group $P3_121$ (a = b = 170.5 Å, c = 161.1 Å) with a solvent content of 70%. Derivatives were prepared by soaking crystals in: 1 mM potassium tetrachloroaurate for 8–12 h, 1 mM mercurous acetate for 6 h, 3 mM methyl mercuric chloride for 6 h, 50 mM samarium acetate for 12 h, or 0.5 mM uranyl acetate for 6–10 h. Diffraction data for native and derivatized crystals were collected at 4 °C at SSRL, beamline 7-1 and beamline 9-1, at 1.08 Å and 0.98 Å respectively. Crystals diffracted reasonably well for about five images using 1° oscillations before showing signs of radiation damage. Taking advantage of the high symmetry and the ability to translate bar shaped crystals allowed for complete data sets to be collected from a few crystals. All data were processed with DENZO and SCALEPACK[29], allowing for decreased resolution and increased mosaicity as the crystal decayed. The CCP4 suite of programs[30] was used for scaling the data sets and locating the heavy atom positions. Patterson maps were viewed with XCONTUR[31]. MLPHARE[30] was used for heavy-atom position refinement and phase calculation. Phases were extended to 3.6 Å and improved by solvent flattening and histogram matching using the program DM[32]. The model was built to 3.6 Å resolution using the program O[33], and then, as more data was collected, extended to 3.3 Å. Maps were improved with phase combination using SIGMAA[34]. Cycles of rebuilding, positional refinement, and simulated annealing using X-PLOR[35] were continued until convergence. Excellent electron density exists throughout the model and 99% of the amino acids were visible. Not surprisingly, the two ill-defined regions of the structure are the site of proteolytic cleavage between the two chains and in a surface loop immediately following His 560 of the translocation domain.

Coordinates. Coordinates have been deposited in the Brookhaven Protein Data Bank (accession number 3bta).

Acknowledgments

We thank M. Evenson, B. Santarsiero, B. Spiller and G. Wedemayer for their help in the project, P. Kuhn and M. Soltis for their assistance at beamlines 7-1 and 9-1 at the Stanford Synchrotron Radiation Laboratory, R. Sweet at beamline X12C of the Brookhaven National Laboratory, T. Earnest at the Advanced Light Source, F. Lebeda for helpful discussions, and T. Umland for providing us with the tetanus toxin binding domain coordinates. Financial support for this research was provided by a NSF pre-doctoral fellowship (D.B.L.) and the US Army Medical Research Institute of Infectious Diseases.

Received 25 June 1998; accepted 7 July 1998.

1. Montecucco, C. & Schiavo, G. Structure and function of tetanus and botulinum neurotoxins. *Quarterly Rev. Biophys.* **28**, 423–472 (1995).
2. Krieglstein, K. G., DasGupta, B. R. & Henschen, A. H. Covalent structure of botulinum neurotoxin type A: location of sulfhydryl groups, and disulfide bridges and identification of C-termini of light and heavy chains. *J. Prot. Chem.* **13**, 49–57 (1994).
3. Simpson, L. L. Kinetic studies on the interaction between botulinum toxin type A and the cholinergic neuromuscular junction. *J.Pharmacol. Exp. Ther.* **212**, 16–21 (1980).
4. Umland, T. C. *et al.* Structure of the receptor binding fragment Hc of tetanus neurotoxin. *Nature Struct. Biol.* **4**, 788–792 (1997).
5. Nishiki, T. *et al.*. The high-affinity binding of *Clostridium botulinum* type B neurotoxin to synaptotagmin II associated with gangliosides GT1b/GD1a. *FEBS Lett.* **378**, 253–257 (1996).
6. Montecucco, C. How do tetanus and botulinum toxins bind to neuronal membranes? *TIBS* **11**, 314–317 (1986).
7. Shapiro, R. E. *et al.* Identification of a ganglioside recognition domain of tetanus toxin using a novel ganglioside photoaffinity ligand. *J. Biol. Chem.* **272**, 30380–30386 (1997).
8. Holm, L. & Sander, C. The FSSP database of structurally aligned protein fold families. *Nucleic Acids Res.* **22**, 3600–3609 (1994).
9. Li, J. D., Carroll, J. & Ellar, D. J. Crystal structure of insecticidal delta-endotoxin from *Bacillus thuringiensis* at 2.5 Å resolution. *Nature* **353**, 815–821 (1991).
10. Choe, S. *et al.* The crystal structure of diphtheria toxin. *Nature* **357**, 216–222 (1992).
11. Parker, M. W., Pattus, F., Tucker, A. D. & Tsernoglou, D. Structure of the membrane-pore-forming fragment of colicin A. *Nature* **337**, 93–96 (1989).
12. Allured, V. S., Collier, R. J., Carroll S. F. & McKay, D. B. Structure of exotoxin A of *Pseudomonas aeruginosa* at 3.0-Angstrom resolution. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 1320–1324 (1986).
13. Muchmore, S. W. *et al.* X-ray and NMR structure of human Bcl-x_L, an inhibitor of programmed cell death. *Nature* **381**, 335–341 (1996).
14. Parker, M. W. & Pattus, F. Rendering a membrane protein soluble in water: a common packing motif in bacterial protein toxins. *Trends Biochem. Sci.* **18**, 391–395 (1993).
15. Wiener, M., Freymann, D., Ghosh, P. & Stroud, R. M. Crystal structure of colicin 1a. *Nature* **385**, 461–464 (1997).
16. Harrison, C. J., Hayer-Hartl, M., Di Liberto, M., Hartl, F. & Kuriyan, J. Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. *Science* **276**, 431–435 (1997).
17. Oblatt-Montal, M., Yamazaki, M., Nelson, R. & Montal, M. Formation of ion channels in lipid bilayers by a peptide with the predicted transmembrane sequence of botulinum neurotoxin A. *Protein Science* **4**, 1490–1497 (1995).
18. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J. & Wiley, D. C. Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **387**, 426–430 (1997).
19. Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**, 37–43 (1994).
20. Fass, D., Harrison, S. C. & Kim, P. S. Retrovirus envelope domain at 1.7 Å resolution. *Nature Struct. Biol* . **3**, 465–469 (1996).
21. Lebeda, F.J. & Olson, M.A. Secondary sturctural predictions for the clostridial neurotoxins. *Proteins: Struct. Funct. Genet.* **20**, 293–300 (1994).
22. Yamasaki, S. *et al.* Synaptobrevin/vesicle-associated membrane protein (VAMP) of *Aplysia californica*: structure and proteolysis by tetanus toxin and botulinal neurotoxins type D and F. *Proc. Natl. Acad. Sci. USA* **91**, 4688–4692 (1994).
23. Woody, M. & DasGupta, B. R. Effect of tetranitromethane on the biological activities of botulinum neurotoxin types A, B, and E. *Mol. Cell. Biochem.* **85**, 159–169 (1989).
24. Morante, S. *et al.* X-ray absorption spectroscopy study of zinc coordination in tetanus neurotoxin, astacin, alkaline protease and thermolysin. *Eur. J. Biochem.* **235**, 606–612 (1996).
25. Schmidt, J. J. & Bostian, K. A. Endoproteinase activity of type A botulinum neurotoxin: substrate requirements and activation by serum albumin. *J. Protein Chem.* **16**, 19–26 (1997).
26. Rossetto, O. *et al.* The metallo-proteinase activity of tetanus and botulinum neurotoxins. *J. Physiology (Paris)* **89**, 43–50 (1995).
27. Schmidt, J. J. & Bostian, K. A. Proteolysis of synthetic peptides by type A botulinum neurotoxin. *J. Protein Chem.* **14**, 703–708 (1995).
28. Stevens, R. C., Evenson, M. L., Tepp, W. & DasGupta, B. R. Crystallization and preliminary x-ray analysis of botulinum neurotoxin type A. *J. Mol. Biol.* **222**, 877–880 (1991).
29. Otwinowski, Z., in *Data Collection and Processing*, (eds Sawyer, L., Isaacs, N. & Bailey, S.) 56–62 (Science and Engineering Research Council, Warrington, UK; 1993).
30. *CCP4: A Suite of Programs for Protein Crystallography* (SERC Daresbury Laboratory, Warrington WA4 4AD, UK; 1979).
31. McRee, D. E. A visual protein crystallographic software system for X11/Xview *J. Mol. Graphics* **10**, 44 (1992).
32. Cowtan, K. Joint CCP4 *ESF-EACBM Newsl. Protein Crystallogr.* **31**, 34 (1994).
33. Jones, T. A., Zou, J.-Y, Cowan, S. W. & Kjelgaard, M. Improved methods for the building of protein models in electron density maps and the location of errors in these. *Acta Crystallogr.* **A47**, 110–119 (1991).
34. Read, R. J. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr.* **42**, 140–149 (1986).
35. Brünger, A.T., *X-PLOR v3.8* (Yale University Press, New Haven, CT; 1996).
36. Kraulis, P. J. MOLSCRIPT- a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
37. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281–296 (1991).

33. Stone JR, Marletta MA: **Synergistic activation of soluble guanylate cyclase by YC-1 and carbon monoxide: implications for the role of cleavage of the iron-histidine bond during activation by nitric oxide.** *Chem Biol* 1998, **5**:255-261.

34. •• Yan S-Z, Huang Z-H, Rao VD, Hurley JH, Tang WJ: **Three discrete regions of mammalian adenylyl cyclase form a site for $G_{s\alpha}$ activation.** *J Biol Chem* 1997, **272**:18849-18854.
Scanning mutagenesis, abetted by molecular modeling, provides a comprehensive structural and functional map of the $G_{s\alpha}$-binding site on adenylyl cyclase (AC). This was subsequently confirmed by the structure of the AC–$G_{s\alpha}$ complex [17••] and by a purely genetic approach [36].

35. Skiba NP, Hamm HE: **How $G_{s\alpha}$ activates adenylyl cyclase.** *Nat Struct Biol* 1998, **5**:88-92.

36. Zimmermann G, Zhou DM, Taussig R: **Genetic selection of mammalian adenylyl cyclases insensitive to stimulation by $G_{s\alpha}$.** *J Biol Chem* 1998, **273**:6968-6975.

37. Parent CA, Devreotes PN: **Constitutively active adenylyl cyclase mutant requires neither G proteins nor cytosolic regulators.** *J Biol Chem* 1996, **271**:18333-18336.

38. Chen JQ, Devivo M, Dingus J, Harry A, Li JR, Sui JL, Carty DJ, Blank JL, Exton JH, Stoffel RH *et al.*: **A region of adenylyl cyclase 2 critical for regulation by G-protein βγ subunits.** *Science* 1995, **268**:1166-1169.

39. Vorherr T, Knopfel L, Hofmann F, Mollner S, Pfeuffer T, Carafoli E: **The calmodulin binding domain of nitric oxide synthase and adenylyl cyclase.** *Biochemistry* 1993, **32**:6081-6088.

40. Wu Z, Wong ST, Storm DR: **Modification of the calcium and calmodulin sensitivity of the type I adenylyl cyclase by mutagenesis of its calmodulin binding domain.** *J Biol Chem* 1993, **268**:23766-23768.

41. • Coudart-Cavalli MP, Sismeiro O, Danchin A: **Bifunctional structure of two adenylyl cyclases from the myxobacterium *Stigmatella aurantiaca*.** *Biochimie* 1997, **79**:757-767.
The authors describe the cloning of a bacterial adenylyl cyclase (AC) that has six transmembrane segments and a single catalytic domain. It presumably functions as a homodimer of 12 transmembrane segments. It suggests a missing link in the probable evolution of the complex mammalian AC from a simpler soluble or single transmembrane homodimeric ancestor and perhaps opens a new avenue for genetic analysis of the function of the transmembrane regions.

Unraveling the structures and modes of action of bacterial toxins

D Borden Lacy* and Raymond C Stevens†

The mechanism by which a soluble protein converts into a protein that spans a membrane remains a central question in understanding the molecular mechanism of toxicity of bacterial protein toxins. Using crystallographic structures of soluble toxins as templates, the past year has seen a number of experiments that are designed to probe the membrane state using other structural methods. In addition, crystallographic information concerning the clostridial neurotoxins has emerged, suggesting a novel mechanism of pore formation and new relationships between toxin binding domains.

Addresses
*Department of Chemistry, University of California at Berkeley, Berkeley, CA 94720, USA;
e-mail: lacy@adrenaline.berkeley.edu
†Department of Chemistry and the Ernest Orlando Berkeley National Laboratory, University of California at Berkeley, Berkeley, CA 94720, USA;
e-mail: stevens@adrenaline.berkeley.edu

Current Opinion in Structural Biology 1998, **8**:778–784

http://biomednet.com/elecref/0959440X00800778

Abbreviations

ATR-FTIR	attenuated total reflection - Fourier transform IR
CFT	channel-forming toxin
EF	edema factor
ER	endoplasmic reticulum
Gb_3	globotriaosylceramide
LF	lethal factor
MTS-ET	methanethiosulfonate ethyltrimethylammonium
OB	oligomer-binding
PA	protective antigen
SLT	shiga-like toxin

Introduction

Bacterial protein toxins act using varied mechanisms to yield pathogenesis. Crystal structures have provided insights into these mechanisms of toxicity and have suggested some unexpected relationships among different toxins. This has enabled the classification of toxins into groups according to their structural features (Table 1, Figure 1). The A–B toxins are characterized by at least two functional domains — the B domain binds the target cell, allowing the internalization of the catalytically active A domain. The AB_5 toxins [1], a subset of the A–B toxins, are characterized by their unique hexameric assembly of five binding subunits (B) per single catalytic subunit (A). Channel formation may be required for the transport of the A subunit, but details as to how this occurs are still emerging. The channel-forming toxins (CFTs) known to date are either α-CFTs, if they are predicted to form pores with a helical hairpin, or β-CFTs, if they insert, or are predicted to insert, to form pores comprising a β barrel [2]. An excellent structural review of these proteins, in addition to the superantigens and some nonbacterial toxins, has recently been published [3].

This review will focus on toxins that penetrate membranes, either causing cell leakage and lysis (the pore formers) or delivering a catalytic functional domain into the cytosol (the A–B toxins) (for recent reviews, see [4,5]). First, it will highlight recent results from experiments that were designed to look at pore formation in each of the three subfamilies of toxins: α-CFT, β-CFT and AB_5. Given the difficulties associated with determining the crystal structures of membrane proteins, a number of different techniques are being used to determine the structures of these molecules in their lipid environments: solid-state NMR spectroscopy; attenuated total reflection - Fourier transform IR (ATR-FTIR) spectroscopy; and the use of site-directed cysteine mutagenesis with fluorescent, channel-blocking and electron spin resonance probes. These techniques are being used to test and refine current models of pore formation. Second, recent work geared towards understanding the interaction between toxin binding domains and extracellular receptors will be described. Previous structural studies identified some common folds in toxin binding domains, but have required co-crystallization with the receptor in order to identify the residues that are associated with specificity. Understanding these interactions has applications in antagonist design. Last, despite the many toxin structures determined in the past 12 years, representatives from the clostridial neurotoxin family (tetanus and botulinum neurotoxin) have, until recently, remained unsolved. The new structures from this family, the tetanus toxin binding domain and botulinum neurotoxin serotype A, will be described in the context of what we know about the structures of other A–B toxins.

β-CFT pore formation

The only toxin for which there is crystallographic structural information resembling the membrane-spanning state is *Staphylococcal aureus* α-hemolysin [6] (Figure 1a). This heptameric protein uses two β strands from each monomer to create a 14-stranded β barrel that spans the membrane. Interestingly, recent studies using atomic force microscopy indicate that α-hemolysin can also form hexameric structures in phospholipid bilayers [7•]. The extent to which α-hemolysin, a pore-forming toxin, serves as a prototype for aerolysin [8] (Figure 1c), another pore-forming toxin, and anthrax (Figure 1b), an A–B toxin with the added function of catalytic subunit delivery, is a subject of recent discussion [9].

Aerolysin – a pore former β-CFT

The receptor-binding event for aerolysin is followed by proteolytic activation, oligomerization to a heptamer, membrane insertion and channel formation. A model of the membrane state of aerolysin was generated by fitting the soluble structure (Figure 1c) into low-resolution electron

Table 1

Structural classification of toxins.

Toxin	Subclass	Channel-forming domain fold	Binding domain fold	References
Colicin A channel domain	Pore former	α-CFT	ND	[44]
Colicin E1 channel domain	Pore former	α-CFT	ND	[45]
Colicin Ia	Pore former	α-CFT	β hairpin and helix	[41]
Colicin N fragment	Pore former	α-CFT	Six-stranded β sheet and one helix	[34••]
δ-endotoxin (Cry) (Figure 1j)	Pore former	α-CFT	Three-sheet prism	[46]
δ-endotoxin (CytB) (Figure 1e)	Pore former	β-CFT		[47]
Proaerolysin	Pore former	β-CFT	APT domain?	[10]
α-hemolysin	Pore former	β-CFT		[6]
Perfringolysin O (Figure 1d)	Pore former	β-CFT		[48]
Anthrax PA	A–B	β-CFT	Immunoglobulin-like	[14]
Cholera toxin (Figure 1l)	A–B	AB_5	OB fold	[49]
Heat-labile enterotoxin LT-1	A–B	AB_5	OB fold	[50]
Heat-labile enterotoxin LT-IIb	A–B	AB_5	OB fold	[51]
Pertussis toxin (Figure 1n)	A–B	AB_5	OB fold and APT domain	[52]
Shiga toxin	A–B	AB_5	OB fold	[53]
SLT-I B oligomer	A–B	AB_5	OB fold	[54]
Diphtheria toxin	A–B	α-CFT	β barrel	[55]
Pseudomonas exotoxin A	A–B	α-CFT		[56]
Tetanus binding domain	A–B	ND	β barrel and β trefoil	[38••]
Botulinum neurotoxin	A–B	Not classified	β barrel and β trefoil	[39••]

ND, not determined.

micrograph images of the membrane-spanning state [10]. This model does not have the bias of assuming that the membrane state will look like that of α-hemolysin and it is steadily being refined with new structural data. For example, the use of dichroic FTIR reveals the orientation of the secondary structures relative to the membrane [11•]. The alignment of the α helices perpendicular to the membrane is especially informative, as the majority of the structure's helices are located in domains 1 and 2. The FTIR data aid the fitting of both these domains into the electron micrograph model. The model of aerolysin inserted into the membrane also indicated that a loop from domain 3 would have to move away from a β sheet in order to permit oligomerization. To test this, mutagenesis was used to place a cysteine on both the loop and the β sheet [12•]. The double mutant formed a disulfide bond (as confirmed by crystallography) and was unable to oligomerize.

Studies of aerolysin using ATR-FTIR and hydrogen–deuterium exchange reveal an increase in β-sheet secondary structure upon proteolytic activation, but relatively minor changes are associated with oligomerization and membrane insertion [11•]. This is in contrast to α-hemolysin, which undergoes an increase in β-sheet content upon membrane insertion [13]. These structural observations may be linked to mechanistic differences in pore formation. Although α-hemolysin forms a nonlytic prepore prior to membrane insertion, no evidence for such a nonlytic oligomer exists for aerolysin [11•].

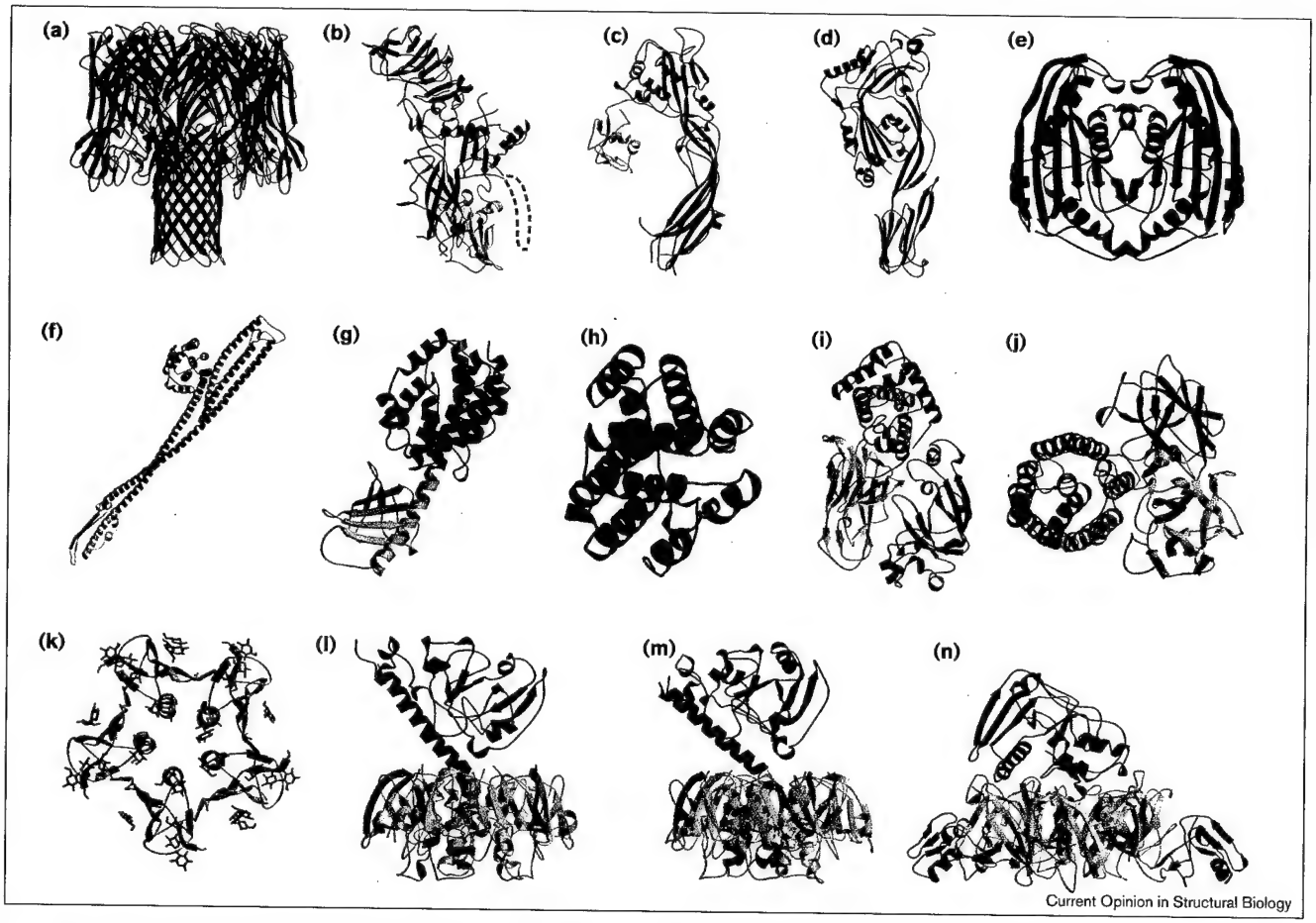
Anthrax – an A–B β-CFT

The B domain of anthrax toxin, protective antigen (PA), interacts with one of two A domains, edema factor (EF) or lethal factor (LF), and mediates their translocation into the cytosol. As the crystal structures of the PA monomer and water-soluble PA heptamer appear to be largely β-sheet structures (Figure 1b), it was suggested that the heptameric β-barrel structure might be responsible for spanning the membrane [14]. In order to test this model for anthrax toxin, whereby the D2L2 loop from seven monomers rearranges to form a 14-stranded β barrel, 24 consecutive residues from this loop were singly substituted with cysteines and derivatized with methanethiosulfonate ethyltrimethylammonium (MTS-ET) [15•]. The effects of these mutations on channel conductance in artificial bilayers were interpreted according to the theory that the positively charged MTS-ET would alter channel conductance in instances when it is located in an ion channel. The observation that channel conductance alternated through two stretches of consecutive residues is consistent with the model whereby two β strands per monomer contribute to channel formation. The observation of multiple conductance step changes supports the notion that an oligomer exists within the membrane.

The existence of an oligomer within the membrane is further supported by recent studies using ATR-FTIR spectroscopy [16•]. The observation of low-frequency absorptions (1623 cm^{-1}) in the characteristic β-strand region indicates that these strands are strongly hydrogen bonded, perhaps due to interoligomer contacts.

If this model is correct, the next issue to be examined is the mechanism by which this β-barrel structure mediates the translocation of EF or LF across the membrane. Further studies using ATR-FTIR indicate that although binding of

Figure 1



Bacterial toxin protein structures. **(a)** *S. aureus* α-hemolysin [6]. The lower 'stalk' in green represents the 14-stranded β barrel, two β strands provided by seven monomers, that spans the membrane. **(b)** *Bacillus anthracis* PA [14]. The dotted line in green represents the D2L2 loop, disordered in the crystal structure, that was singly substituted with cysteine for MTS-ET derivatization [15•]. The receptor-binding domain (yellow) adopts an immunoglobulin-like fold. **(c)** *Aeromonas hydrophilia* proaerolysin [10]. The fold of domain 1 (yellow) has been designated the APT fold and is the putative binding domain. The helices of this fold, as well as the helices (blue) of domain 2, were detected by dichroic FTIR [11•]. **(d)** *Clostridium perfringens* perfringolysin O [48]. **(e)** *B. thuringiensis* mosquitocidal δ-endotoxin CytB [47]. **(f)** *E. coli* colicin Ia [41]. The receptor-binding domain is in yellow, the 160 Å pair of α helices (blue) span the center leading to the translocation domain (blue) and the α-helical pore-forming domain (purple, with the helical hairpin colored green). **(g)** *E. coli* colicin N fragment [34••]. The receptor-binding domain (yellow) is shown in the lower half and the pore-forming domain in the upper half (purple) with the helical hairpin colored green. **(h)** *E. coli* colicin A pore-forming domain [44]. A close-up to show the three layers of ten α helices, with the helical hairpin in green. **(i)** *Corynebacterium diphtheriae* diphtheria toxin [55]. The translocation domain is at the top (purple), with the helical hairpin colored green, the catalytic domain is the lower right half (red) and the receptor-binding domain is the lower left half (yellow), with a jelly-roll topology. **(j)** *B. thuringiensis* insecticidal δ-endotoxin (CryIIIA) [46]. The left half (green) represents the pore-forming domain. The right half is thought to contain the binding activity, with the jelly-roll motif colored blue and the β prism colored yellow. **(k)** *E. coli* shiga-like toxin I B pentamer in complex with a Gb_3 receptor analog [31•]. The structure revealed three Gb_3-binding sites per B monomer. **(l)** *Vibrio cholerae* cholera toxin [49]. The top domain represents the catalytic A subunit, with A_1 colored red and A_2 colored purple. The bottom half (yellow) is an edge on view of the B-subunit pentamer, each monomer representing the OB fold. **(m)** *E. coli* heat-labile enterotoxin [50]. Colored as in (l). **(n)** *Bordetella pertussis* pertussis toxin [52]. This toxin is different from other AB_5 toxins in that the B subunits are not identical. In addition to the OB fold (yellow), the APT domain (blue) is present in two subunits. Figures arranged to emphasize similarities and differences between different classes of toxins. All figures prepared with MOLSCRIPT [57] and Raster3D [58].

the PA to lipids is pH dependent, binding of the EF is not [16•]. The fact that LF [17] and EF [16•] interact with membranes independent of the PA led to the suggestion that the A subunit may play an active role in its own translocation.

α-CFT pore formation in diphtheria and colicin

The pore-forming domains of the α-CFTs are characterized by three layers of 7–10 α helices, in which the putative pore former, the hydrophobic helical hairpin, is shielded in the middle (Figure 1h) [18]. Structural changes with acidic pH are believed to expose this hairpin and allow channel formation. Increasing evidence indicates that these toxins do not insert to form a well-ordered oligomeric structure in the membrane (like the β-CFTs do). For example, in recent experiments using MTS derivatives of 49 cysteine mutants in the helical hairpin of

diphtheria toxin (Figure 1i), multiple step conductances were not observed, suggesting that the channel is not multimeric [19•]. Although this suggestion is supported by results obtained from electron spin resonance spectroscopy, [20], it appears to be a subject of ongoing investigation. Even as a monomer, there are indications that the transmembrane states of the α-CFTs are highly flexible. In the case of colicin A (Figure 1h), previous work indicated a 'molten globule' intermediate for the inserted state [21]. More recent experiments on the colicin E1 channel domain using a number of biophysical techniques — circular dichroism, IR, biosensor, fluorescence and calorimetry — did not identify this intermediate, but did support the idea that the membrane state comprises a two-dimensional, highly helical, flexible array [22•].

A recent paper describing experiments studying the insertion of colicin Ia also contains a discussion on models for the membrane-bound state of colicins in general [23]. These models are supported by recent studies using solid-state NMR spectroscopy, a technique that, like ATR-FTIR spectroscopy, allows the determination of secondary-structure alignments relative to the membrane. A proton-decoupled, ^{15}N one-dimensional spectrum of the colicin B and E1 channel domains indicates that a portion of the peptide bonds, presumably the helical hairpin, aligns perpendicular to the membrane [24•,25•]. Furthermore, two-dimensional, separated local-field spectroscopy (PISEMA) reveals that 50 of the 200 resolved ^{15}N resonances for colicin B within the membrane are in a region that is characteristic of transmembrane helical residues [25•].

Shiga – an AB_5 toxin

The AB_5 toxins are thought to be endocytosed and then translocate their catalytic subunit following retrograde transport to the endoplasmic reticulum (ER) [26]. The process by which this subunit crosses the membrane's lipid bilayer is, however, not well understood. In a recent study using polarized ATR-FTIR, a peptide containing the signal sequence from the A_1 catalytic chain of shiga toxin was characterized as it interacted with the membrane at pH 5 and pH 7 [27•]. Interestingly, the α-helical portion of the peptide was parallel to the membrane at pH 5, but almost perpendicular at pH 7. These results suggest an opposite pH effect compared to the α-CFTs and explains how insertion does not occur at the low pH of the endosome but is then later triggered in the neutral pH environment of the ER lumen. An alternate hypothesis exists that does not involve pore formation. Instead, a conformational change allows the toxin to take advantage of the ER-associated degradation pathway in order to enter the cytosol [28•]. In either case, the mechanism of membrane translocation for these proteins is unclear.

The binding of bacterial toxins to membrane receptors

Although the receptors for many of the bacterial toxins are still unknown, many toxins seem to have an affinity for the sugars that are associated with either glycolipids or glycoproteins on their target membrane. Certain binding domains have similar folds, for example, the oligomer-binding (OB) fold [29] and the APT domain [30]. The identification of common folds has not, however, resulted in an ability to predict the general location of sugar-binding sites. The most recent crystallographic toxin–sugar structure determination, the shiga-like toxin (SLT)-I B pentamer in complex with an analog of its globotriaosylceramide (Gb_3) receptor, differs dramatically from expectation (Figure 1k) [31•]. The structure reveals, surprisingly, three Gb_3-binding sites per B subunit. This can be compared to independent studies carried out in solution using NMR spectroscopy, which could only confirm ganglioside occupancy in 'site 2' [32•]. This apparent discrepancy seems to result from differences in experimental concentrations and reflects the weak binding affinity of each site. This points to an important consideration for toxin–ganglioside binding; that weak binding to many sites could result in a tight apparent binding constant [31•]. The three sites observed in the crystal structure have been incorporated into modeling studies of SLT-IIe in order to explain its ability to bind both Gb_3 and Gb_4 [33•].

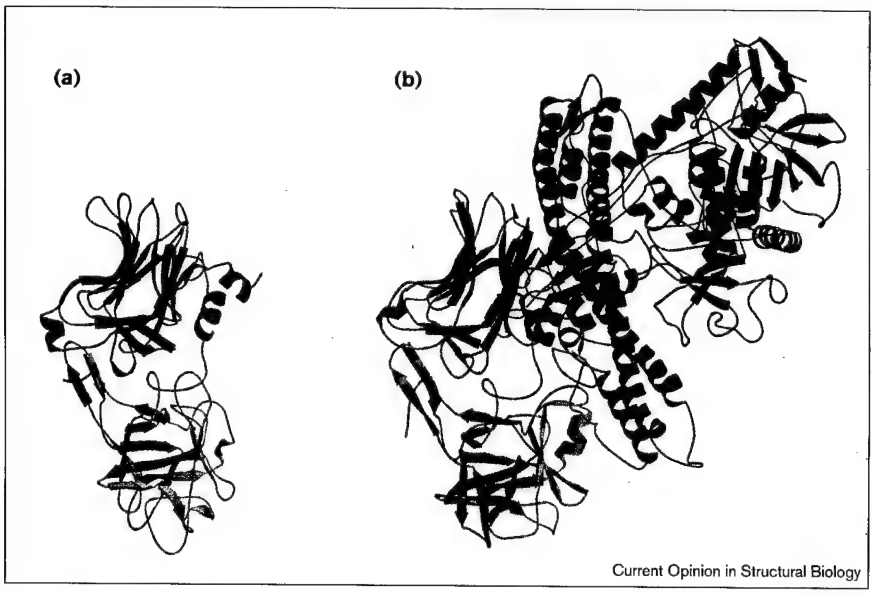
The recent structural elucidation of a colicin N fragment [34••], along with the structure of its ompF receptor [35], makes this one of the best characterized systems for studying colicin binding and translocation. The structure (Figure 1g) reveals a novel binding domain and suggests a model as to how the toxin enters the ompF porin pore. Studies of the binding of colicin N to TolA (a protein involved in translocation across the *Escherichia coli* outer membrane) using isothermal titration microcalorimetry and tryptophan fluorescence indicate that the TolA-binding site is further from the colicin N N terminus than previously thought and imposes new distance constraints on the model of translocation [36•].

Knowing the details of toxin–receptor binding interactions has applications in the design of receptor antagonists. Recent studies of *E. coli* heat-labile enterotoxin (Figure 1m) have identified the binding interactions of four antagonists in the receptor-binding site [37•]. Although the interactions between toxin and receptor seem to be conserved, additional favorable interactions provide an explanation for the increased affinity of the antagonists. The hope is that this can serve as a starting point for increasing the antagonist activity even further.

The clostridial neurotoxins

The crystal structures of both the tetanus toxin binding domain (Figure 2a) [38••] and botulinum neurotoxin type A (Figure 2b) [39••] have recently been reported. The binding domains of both toxins are approximately 50 kDa, structurally homologous and can be divided into two subdomains that are roughly equal in size. The N-terminal subdomain is a β barrel with a jelly-roll topology. This motif has been observed in a number of proteins that

Figure 2



Structures of the clostridial neurotoxins. **(a)** *C. tetani* tetanus neurotoxin receptor-binding domain [38••]. The top subdomain (blue) comprises two seven-stranded β sheets in a jelly-roll topology. An α helix connects it to the C-terminal subdomain (yellow), which adopts a modified β-trefoil fold. **(b)** *C. botulinum* botulinum neurotoxin [39••]. The left third of the molecule corresponds to the receptor-binding domain and is structurally homologous to that of the tetanus neurotoxin, with a mainchain root mean square deviation of 1.5 Å. The middle third, colored in green, represents the translocation domain. Most notable are the 105 Å long pair of α helices and the 'belt' that wraps around the perimeter of the catalytic domain (red) – the right third of the molecule. The catalytic zinc is depicted as a sphere.

interact with sugars [38••,39••]. The C-terminal domain, which contains the putative ganglioside-binding site of both toxins, is a β-trefoil fold. This has been observed in a number of other sugar-binding proteins, including the plant toxins abrin and ricin [40].

It has long been believed that the translocation domain of botulinum neurotoxin resembles that of the α-CFTs, the three-layered array of 7–10 α helices (Figure 1h), as a result of its high helical content as predicted by secondary-structure algorithms. The structure (Figure 2b) reveals an entirely unexpected translocation domain. Most striking is a central pair of 105 Å long α helices that are structurally most similar to the 160 Å long pair of helices observed in colicin IA (Figure 1f) [41]. Unlike colicin IA, which is likely to use this length to span the periplasmic space, there is no anticipated function associated with this length in botulinum neurotoxin. The three layers of α helices observed in the α-CFT channel domains are not seen in the botulinum neurotoxin type A structure. Furthermore, the region that is predicted to be the amphipathic helical pore former [42] is located on the surface and adopts a strand-like conformation. Lastly, the botulinum neurotoxin translocation domain has an approximately 50 residue 'belt' that wraps around the catalytic domain. Although this belt is not necessary for catalytic activity [43], it may play a role in the delivery and release of the catalytic domain into the cytosol. The combination of these unexpected structural features of the translocation domain of botulinum neurotoxin type A suggests that it may employ a heretofore uncharacterized mechanism of pore formation and translocation.

Conclusions

The observation of a few common themes in toxin structure — similar binding domains, channel-forming domains and, even in some cases, catalytic domains — suggests intriguing possibilities for modular toxin assembly from a few structural scaffolds. It provides the toxins with a wide range of cellular targets and effects while requiring a few general mechanisms of entry. The advances of the past year have tested the generalizability of our models for membrane insertion and, with the elucidation of the botulinum neurotoxin type A structure, provided an example of a toxin that is likely to undergo a novel mechanism of pore formation and translocation. The most obvious direction of future study for all of these systems is the determination of high-resolution structures within the membrane. Progress has also been made in the study of toxin binding events, with the complex structures of both SLT-I with receptor and antagonist binding to heat-labile enterotoxin. Knowledge of the nature of the binding interactions will progress as more of the cell-surface receptors are identified.

Acknowledgements

The work in our laboratory on botulinum neurotoxin was carried out in collaboration with the laboratory of Bibhuti R DasGupta and was funded by the US Army Medical Research Institute of Infectious Diseases. We thank Michael Parker and Randy Read for both a critical reading of the manuscript and supplying us with coordinates for the figures. We also thank John Collier for his comments on the manuscript.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Merritt EA, Hol WGJ: **AB_5 toxins**. *Curr Opin Struct Biol* 1995, **5**:165-171.

2. Gouaux E: **Channel-forming toxins: tales of transformation**. *Curr Opin Struct Biol* 1997, **7**:566-573.

3. Parker MW (Ed): *Protein Toxin Structure*. Austin, Texas: RG Landes Co; 1996.

4. Lesieur C, Vecsey-Semjen B, Abrami L, Fivaz M, van der Goot FG: **Membrane insertion: the strategies of toxins.** *Mol Membr Biol* 1997, **14**:45-64.

5. Cabiaux V, Wolff C, Ruysschaert J-M: **Interaction with a lipid membrane: a key step in bacterial toxins virulence.** *Int J Biol Macromol* 1997, **21**:285-298.

6. Song L, Hobaugh MR, Shustak C, Cheley S, Bayley H, Gouaux JE: **Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore.** *Science* 1996, **274**:1859-1866.

7. • Czajkowsky DM, Sheng S, Shao Z: **Staphylococcal α-hemolysin can form hexamers in phospholipid bilayers.** *J Mol Biol* 1998, **276**:325-330.

The observation of a hexamer using atomic force microscopy, when a heptamer was observed using X-ray crystallography, raises the possibility of a polymorphic stoichiometry of subunits existing *in vivo*. The molecular mechanism by which this polymorphism could form and the role of the membrane in its formation need to be further investigated. The observation of incomplete oligomers, if they exist *in vivo*, could imply an alternate mechanism of action, more similar to that of the cholesterol-binding cytolysins.

8. Rossjohn J, Feil SC, McKinstry WJ, Tsernoglou D, van der Goot G, Buckley JT, Parker MW: **Aerolysin – a paradigm for membrane insertion of beta-sheet protein toxins?** *J Struct Biol* 1998, **121**:92-100.

9. Gouaux E: **α-hemolysin from *Staphylococcal aureus*: an archetype of β-barrel, channel-forming toxins.** *J Struct Biol* 1998, **121**:110-122.

10. Parker MW, Buckley JT, Postma JPM, Tucker AD, Leonard K, Pattus F, Tsernoglou D: **Structure of the *Aeromonas* toxin proaerolysin in its water-soluble and membrane-channel states.** *Nature* 1994, **367**:292-295.

11. • Cabiaux V, Buckley JT, Wattiez R, Ruysschaert J-M, Parker MW, van der Goot FG: **Conformational changes in aerolysin during the transition from the water-soluble protoxin to the membrane channel.** *Biochemistry* 1997, **36**:15224-15232.

The authors analyzed the structural changes associated with the proteolytic activation, oligomerization and membrane insertion of aerolysin using attenuated total reflection - Fourier transform IR spectroscopy and hydrogen–deuterium exchange. The results show significant movements upon activation but minor changes upon oligomerization and insertion.

12. • Rossjohn J, Raja SM, Nelson KL, Feil SC, van der Goot FG, Parker MW, Buckley JT: **Movement of loop in domain 3 of aerolysin is required for channel formation.** *Biochemistry* 1998, **37**:741-746.

Crystallography was used to show that an engineered cysteine double mutant of aerolysin forms a disulfide bond between a loop from domain 3 and a β sheet and thus prevents oligomerization. This supports a model whereby the movement of loop 3 precedes and facilitates oligomerization.

13. Vecsey-Semjen B, Lesieur C, Mollby R, van der Goot FG: **Conformational changes due to membrane binding and channel formation by staphylococcal α-toxin.** *J Biol Chem* 1997, **272**:5709-5717.

14. Petosa C, Collier RJ, Klimpel KR, Leppla SH, Liddington RC: **Crystal structure of the anthrax toxin protective antigen.** *Nature* 1997, **385**:833-838.

15. • Benson EL, Huynh PD, Finkelstein A, Collier RJ: **Identification of residues lining the anthrax protective antigen channel.** *Biochemistry* 1998, **37**:3941-3948.

Cysteine substitutions were made for 24 consecutive residues of the anthrax protective antigen D2L2 loop and were derivatized with MTS-ET and tested for changes in channel conductance. An alternating response in two stretches of residues supports a model in which the D2L2 loop from seven monomers rearranges to form a 14-stranded β barrel that spans the membrane.

16. • Wang X-M, Wattiez R, Mock M, Falmagne P, Ruysschaert J-M, Cabiaux V: **Structure and interaction of PA63 and EF (edema toxin) of *Bacillus anthracis* with lipid membrane.** *Biochemistry* 1997, **36**:14906-14913.

Attenuated total reflection - Fourier transform IR spectroscopy and hydrogen–deuterium exchange were used to follow changes in the secondary structures of protective antigen 63, edema factor (EF) and two EF mutants when exposed to lipid and variable pH.

17. Wang X-M, Mock M, Ruysschaert J-M, Cabiaux V: **Secondary structure of anthrax lethal toxin proteins and their interaction with large unilamellar vesicles: a Fourier-transform infrared spectroscopy approach.** *Biochemistry* 1996, **35**:14939-14946.

18. Parker MW, Pattus F: **Rendering a membrane protein soluble in water: a common packing motif in bacterial protein toxins.** *Trends Biochem Sci* 1993, **18**:391-395.

19. • Huynh PD, Cui C, Zhan H, Oh KJ, Collier RJ, Finkelstein A: **Probing the structure of the diphtheria toxin channel.** *J Gen Physiol* 1997, **110**:229-242.

Forty nine of the fifty one residues in the translocation domain helical hairpin of diphtheria toxin were singly substituted with cysteine and derivatized with MTS. Results from channel conductance analysis indicate that the channel is not multimeric and that the helical hairpin may not even be helical within the membrane – suggesting an open, flexible, monomeric structure.

20. Oh KJ, Zhan H, Cui C, Hideg K, Collier RJ, Hubbell WL: **Organization of diphtheria toxin T domain in bilayers: a site-directed spin labeling study.** *Science* 1996, **273**:810-812.

21. van der Goot FG, Gonzalez-Manas JM, Lakey JH, Pattus F: **A 'molten-globule' membrane-insertion intermediate of the pore-forming domain of colicin A.** *Nature* 1991, **354**:408-410.

22. • Zakharov SD, Lindeberg M, Griko Y, Salamon Z, Tollin G, Prendergast FG, Cramer WA: **Membrane-bound state of the colicin E1 channel domain as an extended two-dimensional helical array.** *Proc Natl Acad Sci USA* 1998, **95**:4282-4287.

Numerous biophysical techniques were used to characterize the structural changes within the colicin E1 channel domain upon membrane binding. Results indicate an increase in the helical content but a decrease in the tertiary structure of the membrane state.

23. Kienker PK, Qiu X-Q, Slatin SL, Finkelstein A, Jakes KS: **Transmembrane insertion of the colicin Ia hydrophobic hairpin.** *J Membr Biol* 1997, **157**:27-37.

24. • Kim Y, Valentine K, Opella SJ, Schendel SL, Cramer WA: **Solid-state NMR studies of the membrane-bound closed state of the colicin E1 channel domain in lipid bilayers.** *Protein Sci* 1998, **7**:342-348.

Results from solid-state NMR spectroscopy were found to be consistent with the 38 residue helical hairpin of colicin E1 spanning perpendicular to the membrane.

25. • Lambotte S, Jasperse P, Bechinger B: **Orientational distribution of α-helices in the colicin B and E1 channel domains: a one and two dimensional ^{15}N solid-state NMR investigation in uniaxially aligned phospholipid bilayers.** *Biochemistry* 1998, **37**:16-22.

The alignments of the *trans*-membrane helices of colicin B and E1 were compared at different pHs and different bilayer phospholipid compositions. Results support the 'umbrella' model, whereby the helical hairpin penetrates perpendicular to the membrane.

26. Sandvig K, van Deurs B: **Endocytosis and intracellular sorting of ricin and shiga toxin.** *FEBS Lett* 1994, **346**:99-102.

27. • Menikh A, Saleh MT, Gariepy J, Boggs JM: **Orientation in lipid bilayers of a synthetic peptide representing the C-terminus of the A_1 domain of shiga toxin. A polarized ATR-FTIR study.** *Biochemistry* 1997, **36**:15865-15872.

A synthetic peptide containing a putative signal sequence representing residues 220–246 of the A_1 domain from shiga toxin was shown to convert from a random coil to an α helix in the presence of negatively charged phospholipids. The helix was parallel to the membrane at pH 5, but perpendicular at pH 7. This may suggest a role for the signal sequence in triggering translocation of the A_1 catalytic domain into the cytosol.

28. • Hazes B, Read RJ: **Accumulating evidence suggests that several AB-toxins subvert the endoplasmic reticulum-associated protein degradation pathway to enter target cells.** *Biochemistry* 1997, **36**:11051-11054.

In this model, the A and B subunits of the toxin dissociate within the ER, exposing a hydrophobic surface that 'disguises' the A subunit as a misfolded protein. It enters the ER-associated degradation pathway, but escapes ubiquitin-mediated degradation due to a lack of lysine residues.

29. Murzin AG: **OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences.** *EMBO J* 1993, **12**:861-867.

30. Rossjohn J, Buckley JT, Hazes B, Murzin AG, Read RJ, Parker MW: **Aerolysin and pertussis toxin share a common receptor-binding domain.** *EMBO J* 1997, **16**:3426-3434.

31. • Ling H, Boodhoo A, Hazes B, Cummings MD, Armstrong GD, Brunton JL, Read RJ: **Structure of the shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb_3.** *Biochemistry* 1998, **37**:1777-1788.

The co-crystal structure reveals three Gb_3-binding sites per B subunit of shiga-like toxin I and may explain how tight cell-surface binding could be achieved despite only weak binding to its soluble receptor.

32. • Shimizu H, Field RA, Homans SW, Donohue-Rolfe A: **Solution structure of the complex between the B-subunit homopentamer of verotoxin VT-1 from *Escherichia coli* and the trisaccharide**

moiety of globotriaosylceramide. *Biochemistry* 1998, **37**:11078-11082.

NMR spectroscopy was used to solve the structure of the shiga-like toxin I B pentamer in complex with an analog of its receptor, ^{13}C-enriched Gb_3OSE. A comparison was made to the crystal structure [31•], as the NMR structure only revealed binding at 'site 2'.

33. Cummings MD, Ling H, Armstrong GD, Brunton JL, Read RJ:
• **Modeling the carbohydrate-binding specificity of pig edema toxin.** *Biochemistry* 1998, **37**:1789-1799.

Using the three sites observed in Gb_3 binding to SLT-I [31•], modeling was used to try to explain why SLT-IIe can bind to both Gb_3 and Gb_4, whereas a double mutant (Q65E/K67Q) preferentially binds Gb_3. An atomic understanding of the role of both the additional sugar moiety and the residues involved in binding specificity should provide a basis for therapeutic development.

34. Vetter IR, Parker MW, Tucker AD, Lakey JH, Pattus F, Tsernoglou D:
•• **Crystal structure of a colicin N fragment suggests a model for toxicity.** *Structure* 1998, **6**:863-874.

The structure of a colicin N fragment reveals a helical pore-forming domain, which is similar to those seen in other α-CFTs, and a novel receptor-binding domain. A model for binding and translocation is proposed.

35. Cowan SW, Schirmer T, Rummel G, Steiert M, Ghosh R, Pauptit RA, Jansonius JN, Rosenbusch JP: **Crystal structures explain functional properties of two *E. coli* porins.** *Nature* 1992, **358**:727-733.

36. Raggett EM, Bainbridge G, Evans LJA, Cooper A, Lakey JH:
• **Discovery of critical TolA-binding residues in the bactericidal toxin colicin N: a biophysical approach.** *Mol Microbiol* 1998, **28**:1335-1343.

Binding studies using isothermal titration microcalorimetry and tryptophan fluorescence indicate that the isolated N-terminal domain of colicin N has a higher affinity for TolA than the holotoxin. This result and similarities to bacteriophage DNA transport systems led the authors to suggest that the TolA-binding site is revealed after 'primary receptor binding'. The TolA-binding site appears to be in a different location than previously thought.

37. Merritt EA, Sarfaty S, Feil IK, Hol WGJ: **Structural foundation for the**
• **design of receptor antagonists targeting *Escherichia coli* heat-labile enterotoxin.** *Structure* 1997, **5**:1485-1499.

This paper presents the structure of heat-labile enterotoxin in complex with four receptor antagonists. These compounds represent promising leads as they preserve key toxin–sugar interactions and are considerably smaller than the GM1 receptor. In one case, the receptor antagonist displaces a conserved water molecule, perhaps accounting for the observed gain in affinity.

38. Umland TC, Wingert LM, Swaminathan S, Furey WF, Schmidt JJ,
•• Sax M: **Structure of the receptor binding fragment Hc of tetanus neurotoxin.** *Nat Struct Biol* 1997, **4**:788-791.

This paper describes the structure of the tetanus toxin binding domain and gives a discussion of structures with homologous folds and primary sequence conservation with the botulinum neurotoxin serotypes.

39. Lacy DB, Tepp W, Cohen AC, DasGupta BR, Stevens RC: **The**
•• **crystal structure of botulinum neurotoxin type A and implications for toxicity.** *Nat Struct Biol* 1998, **5**:898-902.

The crystal structure of botulinum neurotoxin type A provided a look at the overall architecture of the catalytic, translocation and binding domains, in addition to discussing how the observed structural features may affect toxicity.

40. Murzin AG, Lesk AM, Chothia C: **β-trefoil fold. Patterns of structure and sequence in the Kunitz inhibitors, interleukins-1β and 1α and fibroblast growth factors.** *J Mol Biol* 1992, **223**:531-543.

41. Wiener M, Freymann D, Ghosh P, Stroud RM: **Crystal structure of colicin 1a.** *Nature* 1997, **385**:461-464.

42. Oblatt-Montal M, Yamazaki M, Nelson R, Montal M: **Formation of ion channels in lipid bilayers by a peptide with the predicted transmembrane sequence of botulinum neurotoxin A.** *Protein Sci* 1995, **4**:1490-1497.

43. Blasi J, Chapman ER, Link E, Binz T, Yamasaki S, De Camilli P, Sudhof TC, Niemann H, Jahn R: **Botulinum neurotoxin A selectively cleaves the synaptic protein SNAP-25.** *Nature* 1993, **365**:160-163.

44. Parker MW, Pattus F, Tucker AD, Tsernoglou D: **Structure of the membrane-pore-forming fragment of colicin A.** *Nature* 1989, **337**:93-96.

45. Elkins P, Bunker A, Cramer WA, Stauffacher CV: **A mechanism for toxin insertion into membranes is suggested by the crystal structure of the channel-forming domain of colicin E1.** *Structure* 1997, **5**:443-458.

46. Li J, Carroll J, Ellar DJ: **Crystal structure of insecticidal δ-endotoxin from *Bacillus thuringiensis* at 2.5 Å resolution.** *Nature* 1991, **353**:815-821.

47. Li J, Koni PA, Ellar DJ: **Structure of the mosquitocidal δ-endotoxin CytB from *Bacillus thuringiensis* sp. Kyushuensis and implications for membrane pore formation.** *J Mol Biol* 1996, **257**:129-152.

48. Rossjohn J, Feil SC, McKinstry WJ, Tweten RK, Parker MW: **Structure of a cholesterol-binding, thiol-activated cytolysin and a model of its membrane form.** *Cell* 1997, **89**:685-692.

49. Zhang RG, Scott DL, Westbrook ML, Nance S, Spangler BD, Shipley GG, Westbrook EM: **The three-dimensional crystal structure of cholera toxin.** *J Mol Biol* 1995, **251**:563-573.

50. Sixma TK, Pronk SE, Kalk KH, Wartna ES, van Zanten BAM, Witholt B, Hol WGJ: **Crystal structure of a cholera toxin-related heat-labile enterotoxin from *E. coli.*** *Nature* 1991, **351**:371-378.

51. van den Akker F, Sarfaty S, Twiddy EM, Connell TD, Holmes RK, Hol WGJ: **Crystal structure of a new heat-labile enterotoxin, LT-IIb.** *Structure* 1996, **4**:665-678.

52. Stein PE, Boodhoo A, Armstrong GD, Cockle SA, Klein MH, Read RJ: **The crystal structure of pertussis toxin.** *Structure* 1994, **2**:45-57.

53. Fraser ME, Chernaia MM, Kozlov YV, James MNG: **Crystal structure of the holotoxin from *Shigella dysenteriae* at 2.5 Å resolution.** *Nat Struct Biol* 1994, **1**:59-64.

54. Stein PE, Boodhoo A, Tyrell GJ, Brunton JL, Read RJ: **Crystal structure of the cell-binding B oligomer of verotoxin-1 from *E. coli.*** *Nature* 1992, **355**:748-750.

55. Choe S, Bennett MJ, Fujii G, Curmi PMG, Kantardjieff KA, Collier RJ, Eisenberg D: **The crystal structure of diphtheria toxin.** *Nature* 1992, **357**:216-222.

56. Allured VS, Collier RJ, Carroll SF, McKay D: **Structure of exotoxin A of *Pseudomonas aeruginosa* at 3.0 Å resolution.** *Proc Natl Acad Sci USA* 1986, **83**:1320-1324.

57. Kraulis PJ: **MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures.** *J Appl Crystallogr* 1991, **24**:946-950.

58. Merritt EA, Murphy MEP: **Raster3D version 2.0. A program for photorealistic molecular graphics.** *Acta Crystallogr D* 1994, **50**:869-873.

Reprinted from *J. Mol. Biol.* (1999) **00**, 000–000

JMB

Sequence Homology and Structural Analysis of the Clostridial Neurotoxins

D. Borden Lacy and Raymond C. Stevens

Article No. jmbi.1999.2945 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (1999) 00, 1–13

JMB

Sequence Homology and Structural Analysis of the Clostridial Neurotoxins

D. Borden Lacy[1] and Raymond C. Stevens[2]*

[1]*Department of Chemistry University of California at Berkeley, Berkeley, CA 94720, USA*

[2]*Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA*

**Corresponding author*

The clostridial neurotoxins (CNTs), comprised of tetanus neurotoxin (TeNT) and the seven serotypes of botulinum neurotoxin (BoNT A-G), specifically bind to neuronal cells and disrupt neurotransmitter release by cleaving proteins involved in synaptic vesicle membrane fusion. In this study, multiple CNT sequences were analyzed within the context of the 1277 residue BoNT/A crystal structure to gain insight into the events of binding, pore formation, translocation, and catalysis that are required for toxicity. A comparison of the TeNT-binding domain structure to that of BoNT/A reveals striking differences in their surface properties. Further, the solvent accessibility of a key tryptophan in the C terminus of the BoNT/A-binding domain refines the location of the ganglioside-binding site. Data collected from a single frozen crystal of BoNT/A are included in this study, revealing slight differences in the binding domain orientation as well as density for a previously unobserved translocation domain loop. This loop and the conservation of charged residues with structural proximity to putative pore-forming sequences lend insight into the CNT mechanism of pore formation and translocation. The sequence analysis of the catalytic domain revealed an area near the active-site likely to account for specificity differences between the CNTs. It revealed also a tertiary structure, highly conserved in primary sequence, which seems critical to catalysis but is 30 Å from the active-site zinc ion. This observation, along with an analysis of the 54 residue "belt" from the translocation domain are discussed with respect to the mechanism of catalysis.

Keywords: clostridial neurotoxin; botulinum neurotoxin; tetanus neurotoxin; translocation; X-ray crystallography

Introduction

The *Clostridium botulinum* organism produces seven immunologically distinct forms of botulinum neurotoxin (BoNT), designated A-G, from different strains (Simpson, 1989). BoNT/E and BoNT/F have been observed from *Clostridium butyricum* and *Clostridium baratti*, respectively (Aureli *et al.*, 1986; Hall *et al.*, 1985). Each toxin is synthesized as an inactive ~150 kDa single-chain protein. The protein is post-translationally proteolyzed to form the active dichain molecule in which the two chains, ~50 and ~100 kDa, remain linked by a disulfide bond (Figure 1(a)). The active dichain molecule is comprised of three ~50 kDa functional domains: binding, translocation, and catalytic (Montecucco & Schiavo, 1995). The binding domain comprises the C-terminal half of the ~100 kDa chain, while the translocation domain is located in its N-terminal half. The catalytic domain, a zinc-endopeptidase, is confined to the N-terminal 50 kDa chain.

These domains correlate with a three-step model for toxicity (Simpson, 1980). In the first step, the binding domain mediates interaction between the toxin and the presynaptic nerve terminal membrane (Dolly *et al.*, 1984). This interaction is thought to occur through both a ganglioside and protein receptor (Montecucco, 1986). Following binding, the protein is internalized by receptor-mediated endocytosis (Black & Dolly, 1986a,b). The second step is triggered by the lower pH of the endosome. Acidic pH is thought to cause a struc-

Abbreviations used: CNT, clostridial neurotoxin; TeNT, tetanus neurotoxin; BoNT, botulinum neurotoxin; VAMP, vesicle-associated membrane protein; SNAP, synaptosomal-associated protein.

E-mail address of the corresponding author: stevens@scripps.edu

0022-2836/99/0000001–13 $30.00/0

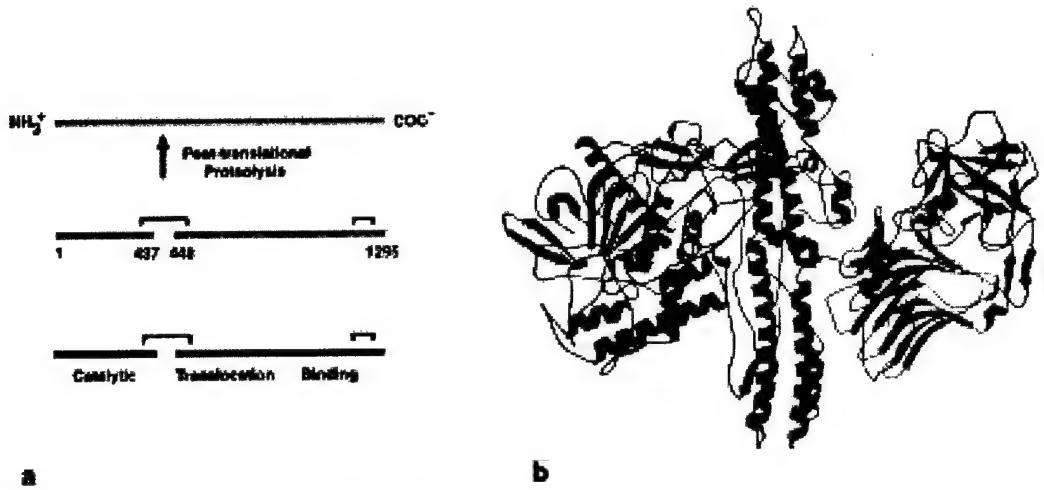


Figure 1. (a) The 150 kDa toxin is post-translationally proteolyzed to form the activated dichain molecule (the arrow in the top part of the Figure illustrates where proteolysis occurs). Two disulfides exist, one that links the two chains and the other in the C-terminal half of the binding domain. The three functional domains are each ~50 kDa and correspond to the catalytic domain (1-437), the translocation domain (448-872), and the binding domain (873-1295), where the numbers refer to the BoNT/A sequence. (b) A backbone trace of the BoNT/A structure with the catalytic domain in purple, the translocation domain in green, the N-terminal binding subdomain in pink, and the C-terminal binding subdomain in blue. The catalytic zinc ion is represented as a sphere with the helix containing the HEXXH zinc-binding motif in red. Helix α22 of the C-terminal binding subdomain is colored orange to show the relative orientation of the putative ganglioside-binding site to the rest of the molecule. The structure represents residues 1-431 and 450-1295, with residues 432-437 and residues 448-449 at the site of proteolytic cleavage being disordered. This Figure was generated using MOLSCRIPT and Raster3D (Kraulis, 1991; Merritt & Bacon, 1997).

tural change in the translocation domain, allowing it to form a pore in the membrane. This pore allows for the translocation of the catalytic domain across this membrane, gaining access to its cytosolic target. The third step involves the cleavage of one of three proteins involved in synaptic vesicle membrane fusion. BoNT/B, BoNT/D, BoNT/F, and BoNT/G cleave distinct sites within a vesicle-associated membrane protein (VAMP, also referred to as synaptobrevin) (Schiavo *et al.*, 1992, 1993a,c, 1994; Yamasaki *et al.*, 1994a,b). BoNT/A and BoNT/E recognize and cleave distinct sites near the C terminus of SNAP-25 (synaptosomal-associated protein of 25 kDa) (Binz *et al.*, 1994; Blasi *et al.*, 1993a; Schiavo *et al.*, 1993a,b), while BoNT/C1 cleaves syntaxin and SNAP-25 (Blasi *et al.*, 1993b; Foran *et al.*, 1996). VAMP, SNAP-25, and syntaxin are collectively termed the SNARE proteins and have been shown to interact in a four-helix coiled-coil in a step thought to precede synaptic vesicle membrane fusion (Sutton *et al.*, 1998). Cleavage of any one of these three proteins prior to the coiled-coil formation disrupts neurotransmitter release at the neuro-muscular junction. The disruption leads to the flaccid paralysis observed in the disease botulism.

TeNT is produced by *Clostridium tetani* but shares ~65% sequence homology and ~35% identity with the BoNT serotypes. Tetanus toxin is a 150 kDa dichain molecule, which is thought to undergo a three-step mechanism of binding, translocation, and catalysis. The key difference is that while it binds at the presynaptic nerve terminal ending, it undergoes retrograde transport up the axon to act in the central nervous system. The molecular determinant of this alternate localization is unknown but is thought to lie in differences in the protein receptor. TeNT cleaves a site in VAMP, identical with that cleaved by BoNT/B (Schiavo *et al.*, 1992). However, given the different localization, the effect of this cleavage is to inhibit the release of inhibitory neurotransmitter and causes the spastic paralysis seen in tetanus poisoning.

The structure of BoNT/A was recently solved to 3.3 Å resolution using data collected from multiple crystals at 4 °C (Lacy *et al.*, 1998). The structure (Figure 1(b)) showed that the binding domain was structurally similar to the TeNT binding domain (Umland *et al.*, 1997), and could be divided into two subdomains, an N-terminal β-barrel and a C-terminal β-trefoil fold. The translocation domain fold is markedly different from the folds observed in other toxins that undergo pore formation and translocation (Lacy & Stevens, 1998). Most notable are a kinked pair of α-helices, 105 Å in length, and a 54 residue "belt" that wraps around the perimeter of the catalytic domain. The translocation domain occludes access to a large, negatively charged cleft leading into the active-site zinc ion of the catalytic domain. The zinc ion appears to be directly coordinated by His222, His226, Glu261 and by a water-mediated coordination through Glu223.

The CNT sequence divergence is substantial when one considers that the family of proteins has maintained striking similarities in a complex series

of functions. In spite of the divergence though, it is reasonable to presume that the BoNT/A structure will be representative for all of the CNTs. Firstly, there is a greater conservation in predicted secondary structure than in the primary structure (Lebeda & Olson, 1994). Secondly, the structure of the binding domain overlays with the tetanus neurotoxin-binding domain with a root-mean-square deviation (rmsd) of 1.5 Å for 363 Cα atoms. Lastly, the functional requirements, especially those of acid-triggered pore formation and translocation, are rigorous and are not likely to tolerate significant structural changes. It is tempting to think that viewing the primary structure differences within the tertiary structure will lend insight into the molecular determinants of differences between the CNTs. The key differences seem to be an ability to discriminate between protein receptor molecules and in the enzyme specificity for one of eight cleavage sites in the three SNARE proteins. An analysis of the sequence conservation can also help identify regions required in the common functions of ganglioside-binding, pore formation, translocation, and catalysis.

Results and Discussion

The binding domain

The structure of the TeNT-binding domain has been solved by independent investigators to 2.7 Å (Umland *et al.*, 1997; PDB code 1af9) and 1.5 Å (PDB accession code 1A8D). The main chains of the structures differ only in the orientation of two surface loops, located in crystal packing interfaces. As both structures were obtained in the same space group, these packing interface differences, along with the temperature and X-ray source differences are likely to account for the differences in resolution. The structure of BoNT/A at 4 °C, and now −170 °C, reveals the orientation of the binding domain and the accessibility of the surface loops in the presence of the translocation domain (in comparison to the TeNT-binding domain structure). The binding domain tilts away from the plane of the catalytic and translocation domains by ~60°. It projects away from the long axis of the translocation domain by ~40° in the structure from 4 °C data and by ~45° in the structure refined into frozen data. This observation is not surprising, given the relatively small interface between the binding and translocation domains. (The binding domain buries ~400 Å^2 of the translocation domain, while the translocation domain buries ~480 Å^2 of the binding domain.) The relative orientation of this domain could vary even more substantially under physiological conditions. The two sub-domains of the binding domain are linked by an α-helix (α21) and create a cleft in their interface (Figure 2(a) and (c)). The C-terminal subdomain buries ~500 Å^2 of the N-terminal subdomain, while the N-terminal subdomain buries ~540 Å^2 of the C-terminal domain. The

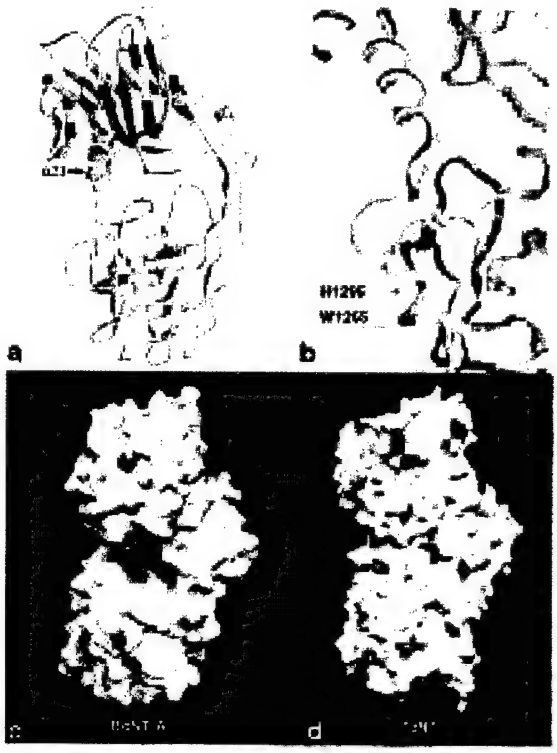


Figure 2. (a) The conserved residues of the BoNT/A-binding domain are colored purple, indicating that the N-terminal sub-domain (top) is more highly conserved than the C-terminal sub-domain (bottom). The second disulfide in BoNT/A, green, is located behind α22. This helix, with residues W1265 and Q1269 colored red, is thought to form part of the ganglioside-binding site. (b) A superposition of the BoNT/A (yellow) and TeNT (red) binding domains at the ganglioside-binding site and a loop difference between the two structures. (c) The two subdomains of the BoNT/A-binding domain are connected by α21 and form an electrostatically positive cleft in their interface. (d) The cleft in TeNT is relatively shallow and neutral, largely due to the loop difference shown in (b).

contacts between the two subdomains are made through loops such that this interface could also be flexible, although the subdomain β-strands align almost identically in the three structures.

The alignment of BoNT/A with TeNT and the six other BoNT serotypes from the three identified organisms is shown along with the secondary structure assignments and solvent accessibility of residues in the BoNT/A structure in Figure 3. There is significantly greater sequence homology within the N-terminal binding subdomain as compared to the C-terminal subdomain (Figures 2(a) and 3). The majority of these highly conserved residues in the N-terminal subdomain, however, point inward, seeming to preserve the hydrophobic core of the β-barrel. As most of the antigenicity conferring serotype uniqueness seems to arise in

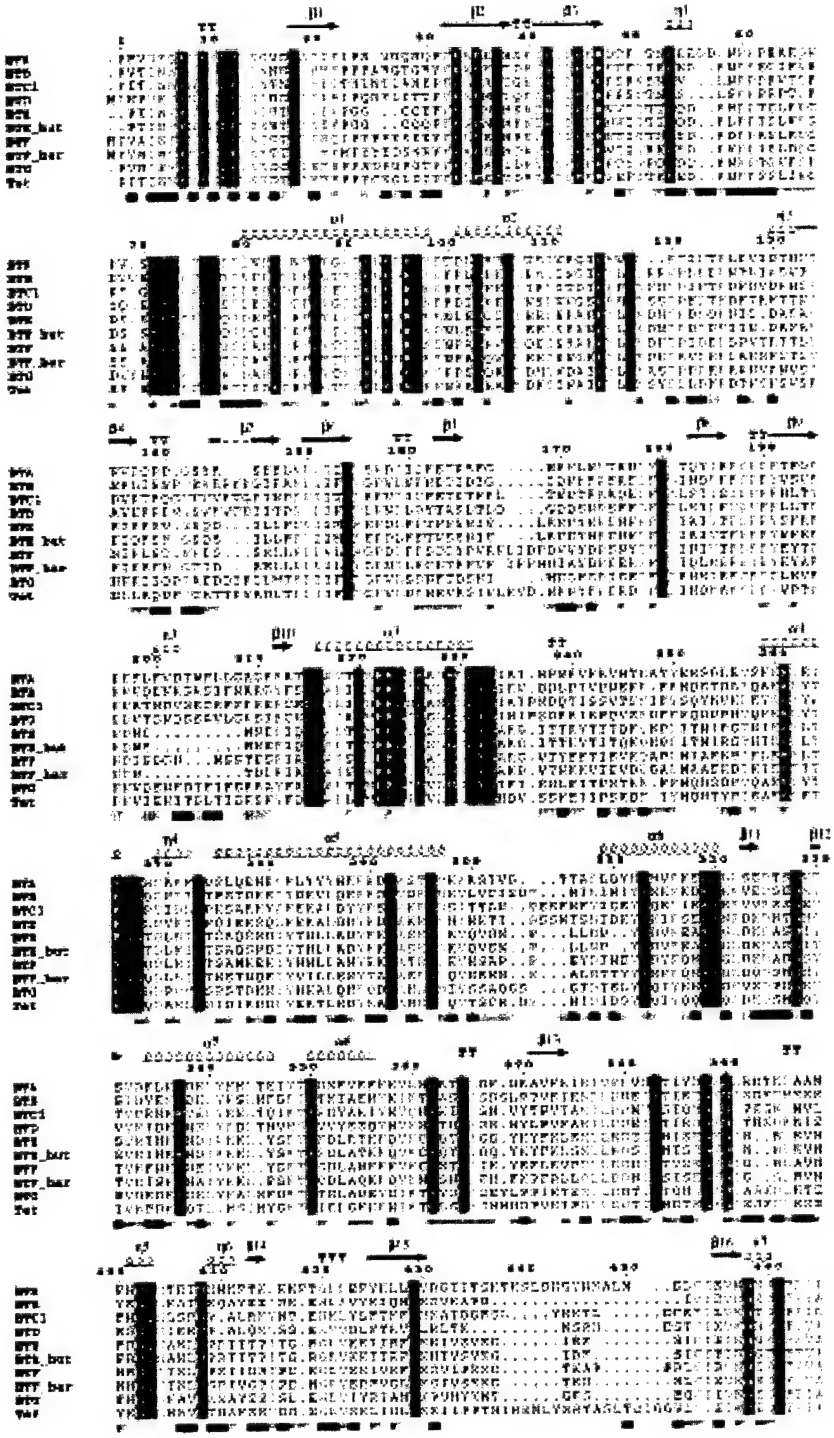


Figure 3 (*legend shown on page 00*)

the binding domain (Chen *et al.*, 1997), it is not surprising that its surface residues vary dramatically among the CNTs.

While the existence of a protein receptor binding site is still under investigation, the ganglioside-binding site is better understood. A binding study using trypsin digests in BoNT/A indicates that the 30 C-terminal residues of the C-terminal subdomain are involved in binding the GT1b ganglioside (Shone *et al.*, 1985). A photoaffinity

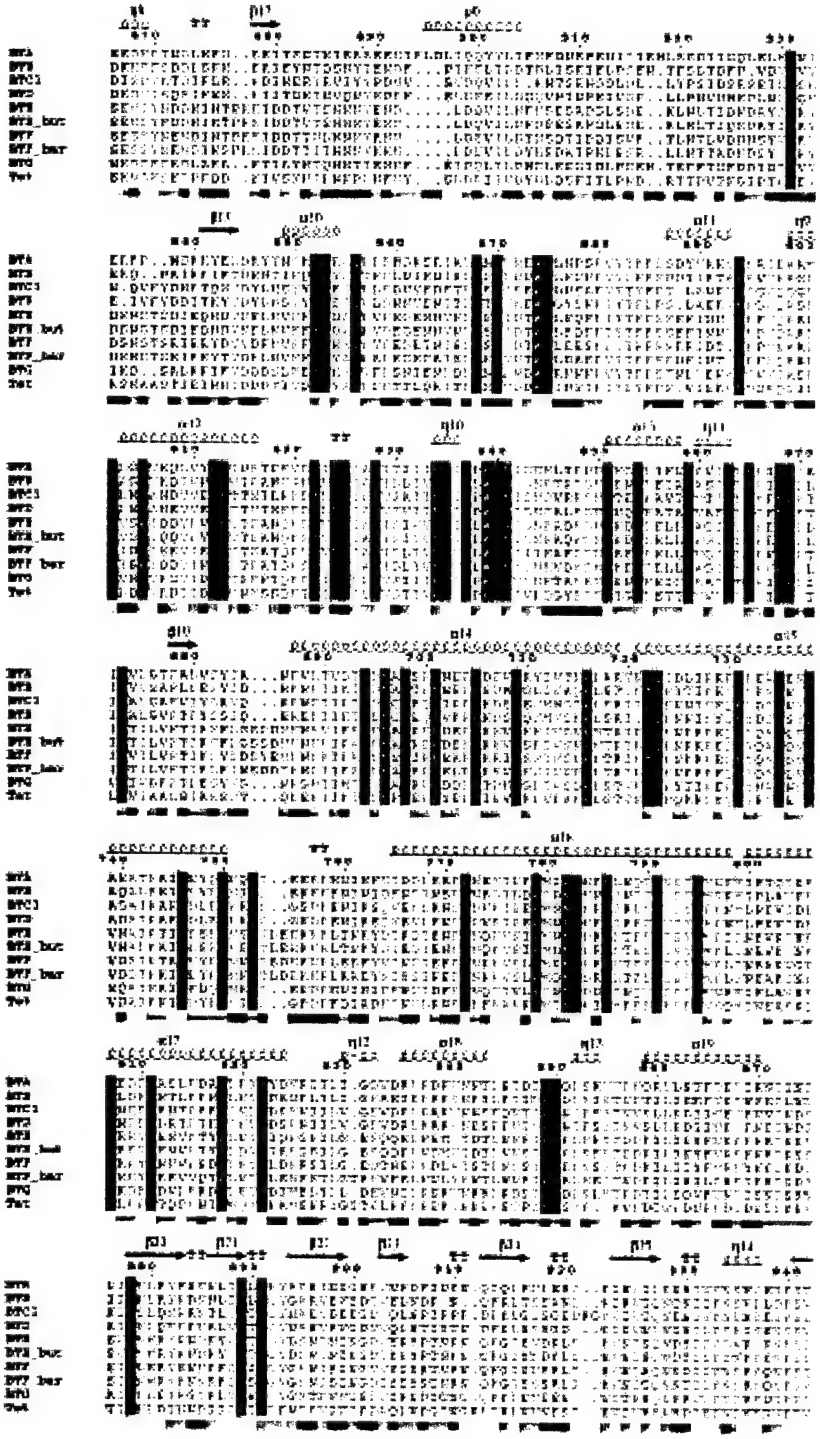


Figure 3 (*legend shown on page 00*)

labeling study using a novel ganglioside probe implicates this same region in TeNT, showing labeling of H1295 (Shapiro *et al.*, 1997). The functional significance of this region has been identified in BoNT/E (Kubota *et al.*, 1997). A monoclonal antibody capable of neutralizing BoNT/E in mice also bound the peptide YLTHMRD. Comparison of this sequence location in BoNT/E aligns to resi-

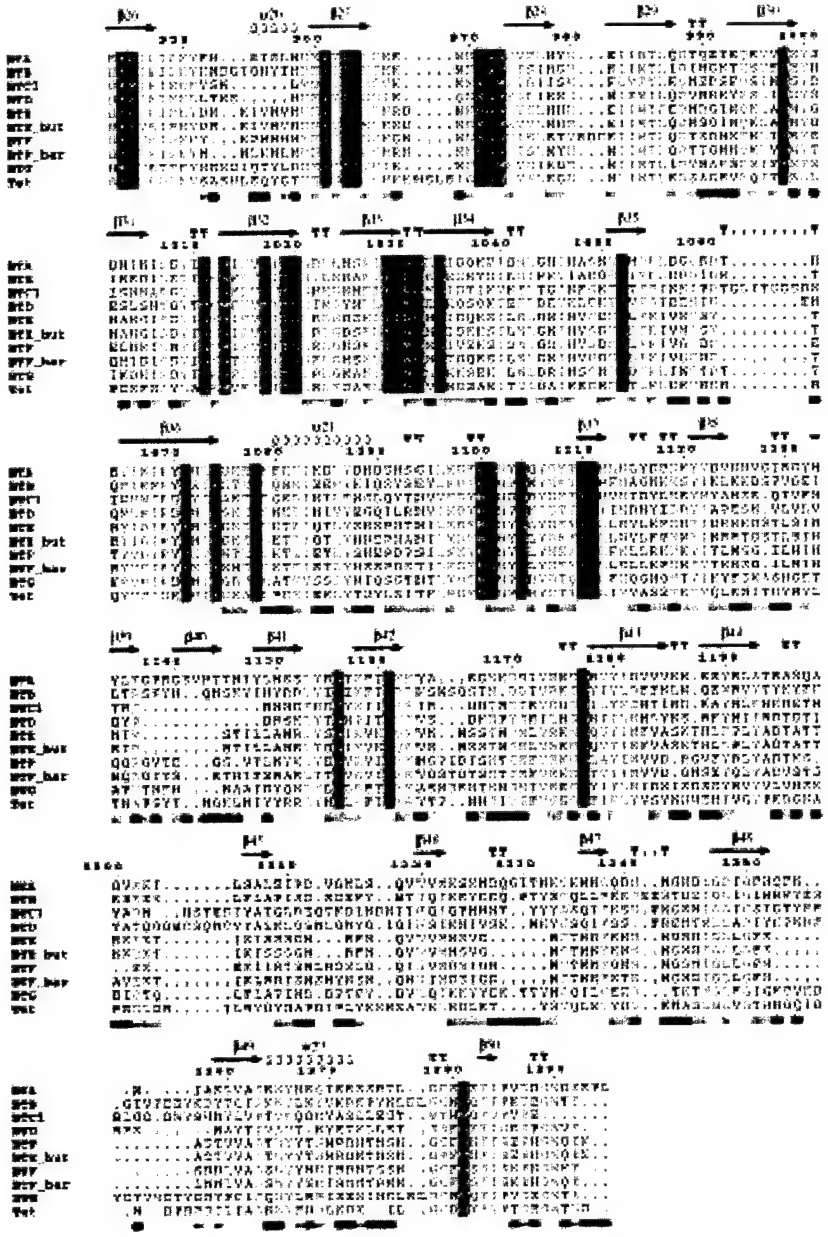


Figure 3 The alignment is displayed to show strict sequence conservation in white letters and a red background, and strong sequence conservation in red letters. The secondary structure elements of the BoNT/A structure are labeled α (α-helix), η (3_{10} helix), β (β-strand) and TT (turn). Assignments were made using DSSP (Kabsch & Sander, 1983) but slightly modified to include helix 20 and strands 31, 40, and 41 after visual inspection. The solvent-accessibility of each residue in the BoNT/A structure is indicated in the bar at the base of the sequences, with white representing buried residues, dark blue representing solvent-accessible residues and light blue representing an intermediate value. This Figure was generated using ESPript (http://www.ipbs.fr/ESPript).

dues 1292-1298 of TeNT and residues 1266-1272 of BoNT/A (Figure 3). Most recently, a binding assay that followed tryptophan fluorescence in BoNT/A showed fluorescence quenching upon binding the ganglioside (Kamata *et al.*, 1997). However, with three tryptophan residues within the 30 C-terminal residues it was impossible to narrow the binding site further. The BoNT/A structure shows that of the three tryptophan residues, only one (W1265) is solvent-accessible (Figure 3). This highly hydrophobic residue is fully exposed to solvent and makes contact with Q1269, the residue in BoNT/A that aligns to H1295 of TeNT, implicated in the photoaffinity labeling study. These two residues are located in the i and $i+4$ positions of helix 22, beneath a loop whose orientation varies dramati-

cally between BoNT/A and TeNT (Figures 1(b) and 2(b)). While in BoNT/A this loop points out, creating a deep positively charged cleft between sub-domains, the loop in TeNT folds in, creating a shallow, more neutral cleft. This surface difference could have a role in the alternative localization of the two toxins.

The translocation domain

The translocation domain is able to form channels in artificial bilayers (Blaustein *et al.*, 1987; Donovan & Middlebrook, 1986; Hoch *et al.*, 1985) and in cell membranes (Sheridan, 1998). Visualization of these channels using electron cryomicroscopy suggests that the channels may be formed by the oligomerization of BoNT to form a tetramer (Schmid *et al.*, 1993). Efforts to identify the pore-forming segment(s) of the CNTs have focussed on identifying amphipathic sequences capable of spanning the membrane (Lebeda & Olson, 1995; Montal *et al.*, 1992). Three such sequences were identified using the MOMENT algorithm for hydrophobic moments (595-614, 625-647, and 648-691) (Lebeda & Olson, 1995). A peptide representing part of one of these sequences, 659-681, was shown to form channels in planar lipid bilayers (Oblatt-Montal *et al.*, 1995). This sequence could possibly oligomerize to form a four-helix bundle in the membrane. The structure of BoNT/A, solved at pH 7, does not refute or support this hypothesis. The previously identified amphipathic sequences do not correspond to the long pairs of kinked α-helices observed in the translocation domain. Instead, the sequences precede these helices and adopt primarily extended loop conformations (Figure 4(a)).

The conservation among the CNTs in the translocation domain is shown by sequence in Figure 3 and structurally in Figure 4(c). While virtually no sequence conservation is observed in the region of the belt, the sequence conservation is fairly high throughout the rest of the translocation domain. The majority of the conserved residues are evenly spaced along the long helices (α14, α15, α16, and α17). The organization seems to preserve the packing between each other and between the smaller helices at either end of the domain (α12 and α18). The only area outside of these helices of particular note is the region from 622-644. This sequence is largely loop, with one 3_{10} helix (η10) and is part of an amphipathic sequence, 625-647, previously identified (Lebeda & Olson, 1995). The loop extends away from the domain pointing toward the binding domain but is not involved in any inter-domain contacts. To our knowledge, this sequence has not been investigated for its ability to form pores in membranes.

The critical follow-up question after identifying the pore-forming segment is to identify the molecular mechanism by which pH triggers this sequence to change structure and form a membrane-spanning channel. The translocation domain is geared to sense this environmental variable, as the calculated pI for its sequence is 4.66 (as compared to 9.1 for the binding domain

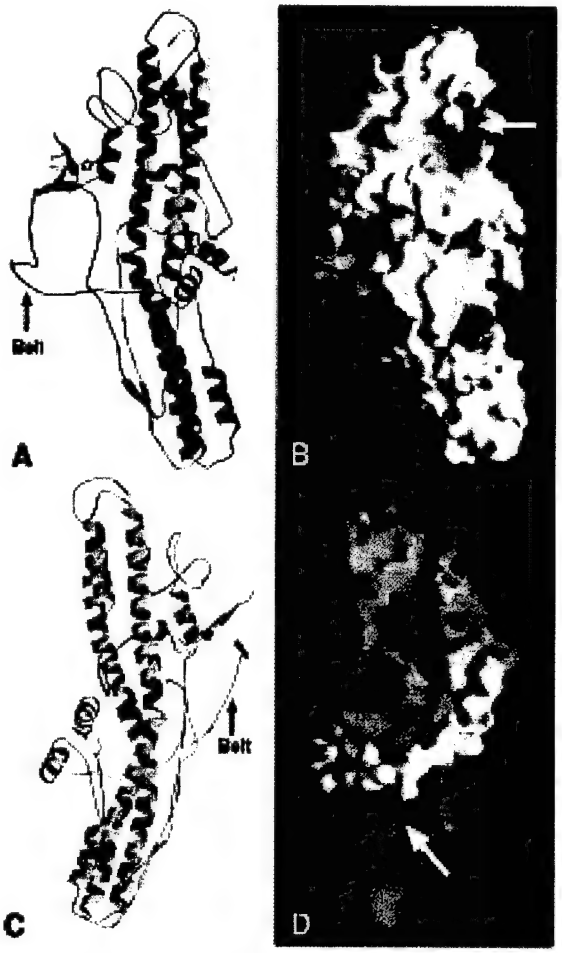


Figure 4. (a) and (b) show identical orientations of the BoNT/A translocation domain and are related to (c) and (d) by a rotation of ~180° around the long axis of the helices. In (c) and (d), the translocation domain belt is protruding out of the plane of the page. In (a), H551 is shown in purple pointing toward the disulfide (yellow) linking the translocation (green) and catalytic (shown partially in blue) domains. The three amphipathic sequences possibly involved in pore formation are colored orange and red, with conserved residues colored red. These sequences form part of a conserved electrostatic cluster shown with an arrow in (b). (c) The overall sequence conservation in the translocation domain structure. The conserved residues, shown in purple, act largely to hold the helical arrangement together, the exception being the conserved residues in the extended loops of the amphipathic sequences, shown in (a). The arrow in (d) points to a cluster of charges conserved at the interface between the main body of the translocation domain and the beginning of the translocation domain belt. (a) and (c) were made using MOLSCRIPT (Kraulis, 1991), while the electrostatic potential surfaces shown in (b) and (d) were made using GRASP (Nicholls *et al.*, 1991).

and 6.0 for the whole toxin). This low value is attributable to the first half of the domain containing the amphipathic loops (548-685, pI 4.6) as opposed to the second half containing the kinked pair of α-helices (686-872, pI 8.5). The three-dimensional charge distribution indicates several clusters of negative charges on both sides of the surface (Figure 4(b) and (d)). One cluster (Figure 4(b)) is comprised of D612, D615, E616, E619, and E588. Of these residues, D612 is strictly conserved, E616 and E588 are largely conserved, and E619 is consistently polar. The clustering of negative charge is expected to raise the pK_a of these residues, making them titratable at endosomal pH (4.5-5.5). We therefore propose that these residues and possibly any histidine residues will effect the structural changes required for pore formation.

The translocation domain contains only two histidine residues (H551 and H560) located in and after α10, respectively, such that they immediately follow the translocation domain belt (492-545) and precede a disordered loop in the structure. While H560 is solvent-accessible, H551 is buried at the interface of the catalytic and translocation domains within 7 Å of their connecting disulfide. The idea that H551 protonation could disrupt this interface is supported by the fact that the loop following H551 is disordered in the crystal structure. The original data for BoNT/A showed high *B*-factors for residues 560-585, with 561-568 completely disordered. The newer data, collected from a single frozen crystal, revealed density in this area, although the *B*-factors for residues placed in this density remain high (Figure 4(a)). This region likely represents an area of inherent flexibility and is proximal both to the negatively charged cluster and the buried histidine residue. This feature could facilitate the structural changes that accompany pore formation as well as the translocation of the unfolded catalytic domain through the pore. The role that the belt might play in translocation is unknown, although with movement in this loop region it is not implausible to consider that the belt could move as well. A second electrostatically charged cluster on the other side of the translocation domain (Figure 4(d)) might also contribute to belt movement. This cluster is formed by both positive and negative residues of α14 (K700, E703, K704, D706, E707, and K710), along with negative charges at the beginning of the belt (D473, E478, E487, and E490). Protonation at this interface could interrupt interactions holding the belt to the main body of the translocation domain thus allowing for the translocation and release of the catalytic domain. While the domains are tethered by a disulfide and the interface between the catalytic and translocation domains (not including the belt) includes ~1190 Å^2 of buried surface area, the non-covalent contacts between the belt and catalytic domain are not substantial. The 54 residue belt buries an additional ~1620 Å^2 of the catalytic domain surface area. This translates to ~30 Å^2 per residue of the belt, considerably lower than the ~200 Å^2 buried by an isoleucine residue in a well-packed protein core. The energy required for removing the belt is probably even lower than that indicated by the buried surface area, since the belt is presumably unstructured in the absence of the catalytic domain, thus adding an entropic cost to their association.

The catalytic domain

Previous alignments have indicated that the catalytic domains share up to 36.5% sequence identity, the exception being the catalytic domains of BoNT/B and TeNT, which share 51.6% sequence identity (Kurazono *et al.*, 1992). All sequences contain a HEXXH sequence motif, typical of many zinc proteases. The catalytic domain sequence alignment also shows the strict conservation of two Asp-Pro bonds, one immediately preceding the HEXXH sequence (D215 and P216) and the other at the N terminus of the molecule (D11 and P12). A third Asp-Pro bond is present in all of the sequences, except BoNT/A, where S74 exists instead of proline. Asp-Pro bonds from BoNT/A have been shown to hydrolyze at pH 5 (DasGupta & Evenson, 1992; DasGupta & Tepp, 1991), a fact that may bear relevance given the low pH environment of the endosome and the need to translocate a 50 kDa protein. It should be noted, however, that a truncation after residue 9 inactivates *in vitro* catalysis (Kurazono *et al.*, 1992).

Other than the zinc protease HEXXH sequence motif, the catalytic domains share no sequence similarity with proteins outside of the CNT family. While a DALI search showed that thermolysin, a well-characterized zinc protease, bore the highest structural identity with the BoNT/A catalytic domain, this similarity was weak with a Z-score of 4.6 for 139 residues (Lacy *et al.*, 1998). Nevertheless, a visual comparison of the two enzymes (Figure 5) is useful, because, like BoNT/A, thermolysin coordinates the active-site zinc ion with two histidine residues (His142 and His146), a glutamate residue (Glu166), and a water-mediated glutamate residue (Glu143) (Colman *et al.*, 1972). Structural similarities include the helix containing the HEXXH sequence (α3) and a four-stranded β-sheet (β3, β6, β7, and β8). This accounts for the structural presentation of the two histidine ligands and the glutamate residue that coordinates the activated water in catalysis. The presentation of the fourth ligand is intriguing, as thermolysin presents the Glu166 ligand with a single helix, while BoNT/A has generated a similar presentation by folding with two smaller helices (α4 and α8) end-to-end. As the two BoNT/A helices point in opposite directions, it is uncertain if maintaining this helical periodicity plays a role in the Glu261 presentation.

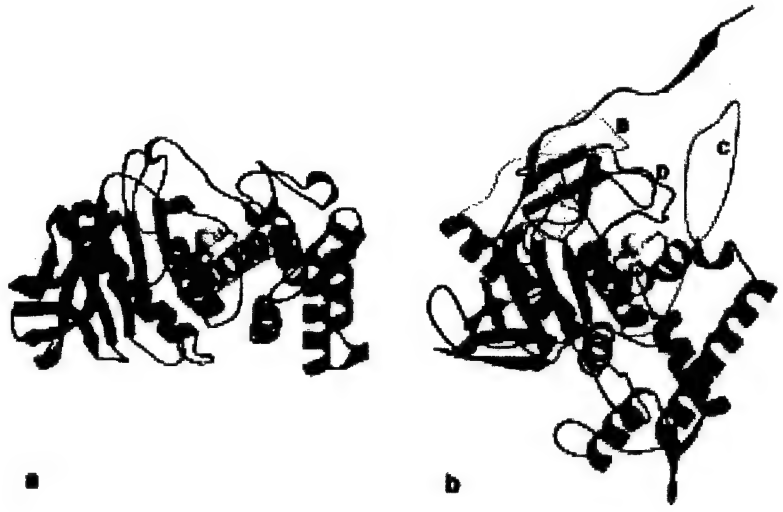


Figure 5. A comparison of (a) thermolysin and (b) the BoNT/A catalytic domain. The helix containing the HEXXH zinc protease motif is colored red, while the helix presenting the second Glu is colored gold. The topology of the secondary structural elements differs such that in BoNT/A, the shorter green and gold helices come together to form the structural equivalent of the single gold helix in thermolysin. Loops B, C, and D of BoNT/A are colored yellow, pink, and orchid, respectively, and form boundaries of accessibility to the active-site zinc ion.

More interestingly, it supports this activity as a likely example of convergent evolution.

The fold comparison highlights the likely cleft by which the substrate gains accessibility to the active-site. This channel of accessibility is bordered by three loops in BoNT/A designated B (47-80), C (231-259), and D (356-371) (Figure 5(b)). The loops in BoNT/A are longer than those in thermolysin and result in a more buried active-site, although these loops could be flexible or oriented differently in the absence of the translocation domain. A second avenue of active-site access would be for substrate to enter from a channel seen arially both in Figure 5(b) and Figure 6(a). The presence of the translocation domain belt and the main body of the translocation domain (Figure 6(b)) occlude both channels of accessibility in the holotoxin.

The conserved residues of the catalytic domain were mapped onto the three-dimensional structure of BoNT/A and are shown in Figure 7. Conserved residues within 6 Å of the active-site cavity are depicted with their full side-chains. In addition to the anticipated residues of the HEXXH motif (H222, E223, and H226), both E260 and E261 (where E261 can coordinate the zinc ion) and R362 and Y365 of loop D are located within this proximity. These seven residues are located in a plane at the base of the active-site channel, thus making it reasonable to assume that their conservation preserves the general zinc endopeptidase activity. The specificity is therefore likely to arise from the residues forming the channel, or cavity, above this base. This cavity is at least 1075 $Å^3$ in the presence of the translocation domain. Non-conserved catalytic domain residues within 4 Å of this cavity include 65-68 of loop B, 161-163, 193, 219, 238 of loop C, 255-258, and 368-9 of loop D. These residues were analyzed in smaller subsets, particularly BoNT/B and TeNT, for sequence conservation, but none was readily apparent. Residues 459, 530, 531, 533, and 536 of the translocation domain also border the active-site cavity where P531 is conserved in all ten sequences and is the only strictly conserved residue of the belt. These residues, which occlude access to the active-site in the holotoxin form, are presumably absent following trans-

Figure 6. A CPK model of (a) the BoNT/A catalytic domain and (b) the BoNT/A holotoxin showing how accessibility to the buried active-site zinc ion (orange) is occluded both by the translocation domain belt and the long axis of the translocation domain (green).

Figure 7. The BoNT/A catalytic domain colored to show residues with high sequence conservation in purple. Seven strictly conserved residues within 6 Å of the active-site zinc ion (gray sphere) are colored orange and include H222, E223, H226, E260, E261, R362, and Y365.

location through the membrane. The catalytic domain may adopt an altered conformation following its translocation, a possibility that should be addressed by the structure determination of this domain in the absence of the neurotoxin.

While some level of sequence conservation was expected in the vicinity of the active-site, the appearance of conservation in two surface loops, distant from the active-site was surprising. Loop A (1-18) is comprised of conserved residues F7, N8, Y9, D11, and P12, and is close to conserved residues of β1, β2, β3, β6, as well as the N terminus of α1 (Figure 8). The importance of this loop was identified in experiments measuring the toxicity of N-terminal and C-terminal catalytic domain deletions (Kurazono *et al.*, 1992). While a deletion of residues 1-7 was tolerated, the deletion of residues 1-9 was not, implicating a direct biochemical and/or indirect structural role for one or both of residues N8 and Y9. While N8 is largely solvent-accessible, Y9 forms part of a conserved charged pocket containing K33 (β2), E46 (β3), D80 (α1), and K83 (α1) (Figure 8). Since D80 and K83 are contained within one end of α1 and the other end of α1 has contact with α3, the HEXXH helix, it is possible that this conservation is merely to preserve the orientation of these helices, and therefore, the structural integrity of the catalytic machinery. This possibility seems unlikely, however, given the 30 Å distance of Y9 from the zinc ion and the lack of sequence conservation in secondary structural elements more proximal to the HEXXH helix (for example, β7 and β8). It seems more likely that the conserved presence of Tyr in the charged pocket is preserving a remote tertiary structure critical for catalysis. This is supported by experiments showing that a monoclonal antibody (mAb) mapping to residues 27-52, when administered intraneurally, neutralized the effects of BoNT/A (Cenci Di Bello *et al.*, 1994). The solvent-accessible residues of this sequence and likely mAb binding site immedi-

Figure 8. The BoNT/A catalytic domain is shown in purple with residues 1-88 colored yellow. Residues 450-550 of the translocation domain are colored green to show the relative position of the translocation domain belt. The Asp-Pro residues are colored red, while the conserved cluster of residues around Tyr9 is shown in purple. The backbone corresponding to the three proline residues P60-P62 is colored green and shows the relative proximity to the active-site zinc ion (gray sphere). Labels are underneath their corresponding secondary structural element.

ately preceed and follow the strands β2 and β3, respectively.

The second conserved surface loop, the last part of loop B, is also located within this proximity. It is intriguing that both loops A and B contain an Asp-Pro pair of residues, although the conservation in loop B did not extend to BoNT/A (Figure 8). While acidic pH is not required for catalysis, cleavage in these sites may aid the efficiency of translocation. The conservation in the latter part of loop B is also of interest, as the preceding residues come into contact with the active-site cavity. An unusual sequence of three consecutive proline residues (P60, P61, and P62) is followed by residues 65-68 (previously noted for their proximity to the channel).

There is mounting evidence suggesting the existence of a remote recognition sequence necessary for catalysis. A ten residue sequence shared by all three of the SNARE proteins has been identified and termed the SNARE motif (Rossetto *et al.*, 1994). This motif is thought to be required for efficient catalysis and is present at varying distances upstream of the cleavage site. A potentially related study shows that the addition of peptides *in trans* upstream and downstream of the TeNT cleavage site is able to activate the toxin (Cornille *et al.*, 1997). The idea of exosite-dependent cleavage implies that a conformational change, in addition to simple removal of the translocation domain, may be required for catalysis. Perhaps the regions in BoNT/A that share high degrees of sequence conservation distal from the active-site represent such general activation domains. Alternatively, these conserved surface loops could represent oligomeric interfaces or be involved in localization of the toxin within the cytosol.

Materials and Methods

Sequence alignments and superposition on the BoNT/A structure

Nine CNT sequences were aligned to that of BoNT/A using CLUSTAL W (Thompson *et al.*, 1994). A gap opening penalty of 10, gap extension penalty of 0.05, and gap separation distance of 8 were used with the BLOSUM62 matrix (Henikoff & Henikoff, 1996). This matrix was also used in the display in Figure 3, using a global score of 0.15. The Risler matrix was used to generate per residue scores for the alignment (Risler *et al.*, 1988). These scores were output in the *B*-factor column of the BoNT/A coordinates, facilitating the display of sequence conservation on the three-dimensional structure. Residues with scores

68 and higher were colored purple, while those below 68 were colored yellow. The number 68 was empirically determined to best match the display of sequence conservation seen in Figure 3.

Toxin structure at −170°C

The protein was concentrated to 5 mg/ml in 10 mM Hepes (pH 7), 100 mM KCl. This stock was mixed 1:1 (v/v) with a mother liquor of 30% (v/v) PEG 600 and 100 mM Tris (pH 7) and allowed to crystallize by hanging drop vapor diffusion. The crystal formed overnight and grew to a size of about 100 μm × 200 μm × 520 μm over the course of a week. The crystal was washed briefly in a solution made from 2 μl of mother liquor and 2 μl of 30% PEG 600 and plunged into liquid nitrogen. Data were collected at SSRL beamline 9-1 and processed with DENZO, yielding data 99.4% complete to 3.3 Å with 98.5% completeness in the last bin. The data were 5.8-fold redundant with an R_{merge} of 8.5% overall. After orienting the molecule with AMoRe, the structure was readily refined into the frozen data using CNS, yielding a final R_{work} of 26% and R_{free} of 31%. While the processing and refinement statistics indicate that the quality is lower than the structure solved at 4°C, these data did bring in density for residues 561-568.

References

Aureli, P., Fenicia, L., Pasolini, B., Gianfranceschi, M., McCroskey, L. M. & Hatheway, C. L. (1986). Two cases of type E infant botulism caused by neurotoxigenic *Clostridium butyricum* in Italy. *J. Infect. Dis.* **154**, 207-211.

Binz, T., Blasi, J., Yamasaki, S., Baumeister, A., Link, E., Sudhof, T. C., Jahn, R. & Niemann, H. (1994). Proteolysis of SNAP-25 by types E and A botulinal neurotoxins. *J. Biol. Chem.* **269**, 1617-1620.

Black, J. D. & Dolly, J. O. (1986a). Interaction of ^{125}I-labeled botulinum neurotoxins with nerve terminals. I. Ultrastructural autoradiographic localization and quantitation of distinct membrane acceptors for types A and B on motor nerves. *J. Cell. Biol.* **103**, 521-534.

Black, J. D. & Dolly, J. O. (1986b). Interaction of ^{125}I-labeled botulinum neurotoxins with nerve terminals. II. Autoradiographic evidence for its uptake into motor nerves by acceptor-mediated endocytosis. *J. Cell. Biol.* **103**, 535-544.

Blasi, J., Chapman, E. R., Link, E., Binz, T., Yamasaki, S., De Camilli, P., Sudhof, T. C., Niemann, H. & Jahn, R. (1993a). Botulinum neurotoxin A selectively cleaves the synaptic protein SNAP-25. *Nature*, **365**, 160-163.

Blasi, J., Chapman, E. R., Yamasaki, S., Binz, T., Niemann, H. & Jahn, R. (1993b). Botulinum neurotoxin C1 blocks neurotransmitter release by means of cleaving HPC-1/syntaxin. *EMBO J.* **12**, 4821-4828.

Blaustein, R. O., Germann, W. J., Finkelstein, A. & DasGupta, B. R. (1987). The N-terminal half of the heavy chain of botulinum type A neurotoxin forms channels in planar phospholipid bilayers. *FEBS Letters*, **226**, 115-120.

Cenci Di Bello, I., Poulain, B., Shone, C. C., Tauc, L. & Dolly, J. O. (1994). Antagonism of the intracellular action of botulinum neurotoxin type A with monoclonal antibodies that map to light-chain epitopes. *Eur. J. Biochem.* **219**, 161-169.

Chen, F., Kuziemko, G. M., Amersdorfer, P., Wong, C., Marks, J. D. & Stevens, R. C. (1997). Antibody mapping to domains of botulinum neurotoxin serotype A in the complexed and uncomplexed forms. *Infect. Immun.* **65**, 1626-1630.

Colman, P. M., Jansonius, J. N. & Matthews, B. W. (1972). The structure of thermolysin: an electron density map at 2-3 Å resolution. *J. Mol. Biol.* **70**, 701-724.

Cornille, F., Martin, L., Lenoir, C., Cussac, D., Roques, B. P. & Fournie-Zaluski, M. C. (1997). Cooperative exosite-dependent cleavage of synaptobrevin by tetanus toxin light chain. *J. Biol. Chem.* **272**, 3459-3464.

DasGupta, B. R. & Evenson, M. L. (1992). Botulinum neurotoxin type A: the isolated light chain breaks down. *FASEB J.* **6**, A227.

DasGupta, B. R. & Tepp, W. (1991). Botulinum neurotoxin type A breaks down at pH 5.0 at the two Asp-Pro bonds of the heavy chain. *Soc. Neurosci. Abstr.* **17**, 1526.

Dolly, J. O., Black, J., Williams, R. S. & Melling, J. (1984). Acceptors for botulinum neurotoxin reside on motor nerve terminals and mediate its internalization. *Nature*, **307**, 457-460.

Donovan, J. J. & Middlebrook, J. L. (1986). Ion-conducting channels produced by botulinum toxin in planar lipid membranes. *Biochemistry*, **25**, 2872-2876.

Foran, P., Lawrence, G. W., Shone, C. C., Foster, K. A. & Dolly, J. O. (1996). Botulinum neurotoxin C1 cleaves both syntaxin and SNAP-25 in intact and permeabilized chromaffin cells: correlation with its blockade of catecholamine release. *Biochemistry*, **35**, 2630-2636.

Hall, J. D., McCroskey, L. M., Pincomb, B. J. & Hatheway, C. L. (1985). Isolation of an organism resembling *Clostridium barati* which produces type F botulinal toxin from an infant with botulism. *J. Clin. Microbiol.* **21**, 654-655.

Henikoff, J. G. & Henikoff, S. (1996). Blocks database and its applications. *Methods Enzymol.* **266**, 88-105.

Hoch, D. H., Romero-Mira, M., Ehrlich, B. E., Finkelstein, A., DasGupta, B. R. & Simpson, L. L. (1985). Channels formed by botulinum, tetanus, and diphtheria toxins in planar lipid bilayers: relevance to translocation of proteins across membranes. *Proc. Natl Acad. Sci. USA*, **82**, 1692-1696.

Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577-2637.

Kamata, Y., Yoshimoto, M. & Kozaki, S. (1997). Interaction between botulinum neurotoxin type A and ganglioside: ganglioside inactivates the neurotoxin and quenches its tryptophan fluorescence. *Toxicon*, **35**, 1337-1340.

Kraulis, P. J. (1991). MOLSCRIPT- a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.

Kubota, T., Watanabe, T., Yokosawa, N., Tsuzuki, K., Indoh, T., Moriishi, K., Sanda, K., Maki, Y., Inoue, K. & Fujii, N. (1997). Epitope regions in the heavy chain of *Clostridium botulinum* type E neurotoxin recognized by monoclonal antibodies. *Appl. Environ. Microbiol.* **63**, 1214-1218.

Kurazono, H., Mochida, S., Binz, T., Eisel, U., Quanz, M., Grebenstein, O., Wernars, K., Poulain, B., Tauc,

L. & Niemann, H. (1992). Minimal essential domains specifying toxicity of the light chains of tetanus toxin and botulinum neurotoxin type A. *J. Biol. Chem.* **267**, 14721-14729.

Lacy, D. B. & Stevens, R. C. (1998). Unraveling the structures and modes of action of bacterial toxins. *Curr. Opin. Struct. Biol.* **8**, 778-784.

Lacy, D. B., Tepp, W., Cohen, A. C., DasGupta, B. R. & Stevens, R. C. (1998). Crystal structure of botulinum neurotoxin type A and implications for toxicity. *Nature Struct. Biol.* **5**, 898-902.

Lebeda, F. J. & Olson, M. A. (1994). Secondary structural predictions for the clostridial neurotoxins. *Proteins: Struct. Funct. Genet.* **20**, 293-300.

Lebeda, F. J. & Olson, M. A. (1995). Structural predictions of the channel-forming region of botulinum neurotoxin heavy chain. *Toxicon*, **33**, 559-567.

Merritt, E. A. & Bacon, D. J. (1997). Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505-524.

Montal, M. S., Blewitt, R., Tomich, J. M. & Montal, M. (1992). Identification of an ion channel-forming motif in the primary structure of tetanus and botulinum neurotoxins. *FEBS Letters*, **313**, 12-18.

Montecucco, C. (1986). How do tetanus and botulinum toxins bind to neuronal membranes? *Trends Biochem. Sci.* **11**, 315-317.

Montecucco, C. & Schiavo, G. (1995). Structure and function of tetanus and botulinum neurotoxins. *Quart. Rev. Biophys.* **28**, 423-472.

Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281-296.

Oblatt-Montal, M., Yamazaki, M., Nelson, R. & Montal, M. (1995). Formation of ion channels in lipid bilayers by a peptide with the predicted transmembrane sequence of botulinum neurotoxin A. *Protein Sci.* **4**, 1490-1497.

Risler, J. L., Delorme, M. O., Delacroix, H. & Henaut, A. (1988). Amino acid substitutions in structurally related proteins. A pattern recognition approach. Determination of a new and efficient scoring matrix. *J. Mol. Biol.* **204**, 1019-1029.

Rossetto, O., Schiavo, G., Montecucco, C., Poulain, B., Deloye, F., Lozzi, L. & Shone, C. C. (1994). SNARE motif and neurotoxins. *Nature*, **372**, 415-416.

Schiavo, G., Benfenati, F., Poulain, B., Rossetto, O., Polverino de Laureto, P., DasGupta, B. R. & Montecucco, C. (1992). Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolytic cleavage of synaptobrevin. *Nature*, **359**, 832-835.

Schiavo, G., Rossetto, O., Catsicas, S., Polverino de Laureto, P., DasGupta, B. R., Benfenati, F. & Montecucco, C. (1993a). Identification of the nerve terminal targets of botulinum neurotoxin serotypes A, D, and E. *J. Biol. Chem.* **268**, 23784-23787.

Schiavo, G., Santucci, A., Dasgupta, B. R., Mehta, P. P., Jontes, J., Benfenati, F., Wilson, M. C. & Montecucco, C. (1993b). Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct COOH-terminal peptide bonds. *FEBS Letters*, **335**, 99-103.

Schiavo, G., Shone, C. C., Rossetto, O., Alexander, F. C. & Montecucco, C. (1993c). Botulinum neurotoxin serotype F is a zinc endopeptidase specific for VAMP/synaptobrevin. *J. Biol. Chem.* **268**, 11516-11519.

Schiavo, G., Malizio, C., Trimble, W. S., Polverino de Laureto, P., Milan, G., Sugiyama, H., Johnson, E. A. & Montecucco, C. (1994). Botulinum G neurotoxin cleaves VAMP/synaptobrevin at a single Ala-Ala peptide bond. *J. Biol. Chem.* **269**, 20213-20216.

Schmid, M. F., Robinson, J. P. & DasGupta, B. R. (1993). Direct visualization of botulinum neurotoxin-induced channels in phospholipid vesicles. *Nature*, **364**, 827-830.

Shapiro, R. E., Specht, C. D., Collins, B. E., Woods, A. S., Cotter, R. J. & Schnaar, R. L. (1997). Identification of a ganglioside recognition domain of tetanus toxin using a novel ganglioside photoaffinity ligand. *J. Biol. Chem.* **272**, 30380-30386.

Sheridan, R. E. (1998). Gating and permeability of ion channels produced by botulinum toxin types A and E in PC12 cell membranes. *Toxicon*, **36**, 703-717.

Shone, C. C., Hambleton, P. & Melling, J. (1985). Inactivation of *Clostridium botulinum* type A neurotoxin by trypsin and purification of two tryptic fragments. Proteolytic action near the COOH- terminus of the heavy subunit destroys toxin-binding activity. *Eur. J. Biochem.* **151**, 75-82.

Simpson, L. L. (1980). Kinetic studies on the interaction between botulinum toxin type A and the cholinergic neuromuscular junction. *J. Pharmacol. Expt. Ther.* **212**, 16-21.

Simpson, L. L. (1989). *Botulinum Neurotoxin and Tetanus Toxin*, Academic Press, San Diego.

Sutton, R. B., Fasshauer, D., Jahn, R. & Brunger, A. T. (1998). Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature*, **395**, 347-353.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673-4680.

Umland, T. C., Wingert, L. M., Swaminathan, S., Furey, W. F., Schmidt, J. J. & Sax, M. (1997). Structure of the receptor binding fragment HC of tetanus neurotoxin. *Nature Struct. Biol.* **4**, 788-792.

Yamasaki, S., Binz, T., Hayashi, T., Szabo, E., Yamasaki, N., Eklund, M., Jahn, R. & Niemann, H. (1994a). Botulinum neurotoxin type G proteolyses the Ala81-Ala82 bond of rat synaptobrevin 2. *Biochem. Biophys. Res. Commun.* **200**, 829-835.

Yamasaki, S., Hu, Y., Binz, T., Kalkuhl, A., Kurazono, H., Tamura, T., Jahn, R., Kandel, E. & Niemann, H. (1994b). Synaptobrevin/vesicle-associated membrane protein (VAMP) of *Aplysia californica*: structure and proteolysis by tetanus toxin and botulinal neurotoxins type D and F. *Proc. Natl Acad. Sci. USA*, **91**, 4688-4692.

Edited by G. Von Heijne

(Received 29 March 1999; received in revised form 8 June 1999; accepted 9 June 1999)

Structure
Submitted, 1999

Three-dimensional reconstruction of Botulinum Neurotoxin Complex, Serotype A

Key Words: electron microscopy, three-dimensional reconstruction, botulinum neurotoxin complex

Flora Chen[1], Geoffrey M. Kuziemko[2], Raymond C. Stevens[2]*

[1]Graduate Group in Biophysics, University of California, Berkeley, California 94720

[2]Department of Chemistry, University of California, Berkeley, CA 94720

*Corresponding Author
E-mail: stevens@adrenaline.berkeley.edu
Phone (510) 643-8285
Fax (510) 643-0883

Background: The botulinum neurotoxin complex protects the botulinum neurotoxin from pH extremes and proteolysis, increasing the neurotoxin's oral toxicity. While the x-ray crystal structure of botulinum neurotoxin has recently been determined, the structure of the botulinum neurotoxin complex remains unknown. To determine the structure of the complex, we used low dose electron microscopy on a tilt-series of two-dimensional protein crystals.

Results: A three-dimensional structure of the complex was reconstructed from 109 images with tilt angles ranging from -60° to +60°. The structure of the complex occupies a unit cell with dimensions of a = 157(±2) Å, b = 159(±4) Å and γ = 120°. Inside the unit cell is a triangular ring of densities (~112 Å long x 90 Å high) surrounded by three smaller globular densities (~67 Å long x 78 Å high). The approximate calculated mass of the botulinum complex monomer was 500 kD.

Conclusions: Because the typical mass of the botulinum neurotoxin complex from Serotype A is 900 kD, the Serotype A complex is likely a dimer of two 500 kD complexes, with each 500 kD complex containing a single molecule of botulinum neurotoxin. Inside each 500 kD complex, the botulinum neurotoxin is surrounded by a non-toxic, non-hemagglutinating protein and a hemagglutinin. The presence of protein "armor" surrounding the neurotoxin likely provides the protection for the neurotoxin to withstand the destructive conditions of the gastrointestinal tract.

Introduction

The anaerobic bacteria *Clostridium botulinum* produce seven serotypes (A-G) of botulinum neurotoxins, which cause the deadly illness, botulism. Due to their specificities for neural synapses and their enzymatic activity, botulinum neurotoxins are the most toxic substances known to man[1]. Despite their toxicities, the botulinum neurotoxin molecules are highly susceptible to degradation at extremes of pH, temperature and proteolytic cleavage [2]; [3]. How does such an unstable molecule act as such a potent oral toxin? A comparison of botulinum neurotoxin with tetanus toxin shows that although both toxins share many

characteristics—molecular weight, similar macrostructure, similar mechanism of action, and a strong amino acid sequence homology—only botulinum neurotoxin is a food poison[4]. Researchers believe that a complex of proteins that is produced by *Clostridium botulinum*, and is known to associate with the neurotoxin, produces the high oral toxicity of botulinum neurotoxin [5]; [4]; [2]. Produced in food or in culture, the neurotoxin is always complexed with at least one non-toxic protein[6]. The complexes of serotype A exist in three size —300 kD, 500 kD and 900 kD—respectively known as 12S, 16S and 19S from ultracentrifugation measurements [7]. The larger complexes have hemagglutinating as well as toxic ability [8]. Although the botulinum neurotoxin complexes act as potent oral toxins, much research has focused only on the neurotoxic portion. Consequently, very little structural information is known of the whole complex.

Previous attempts to analyze the complex using electron microscopy failed to visualize the entire complex and only weakly visualized the toxic and non-toxic components [9]. Some success visualizing the structure of botulinum neurotoxin Serotype B has been achieved, but a structure of the complex remains unrealized [10]; [11]. Lipid-layer two-dimensional crystallization technique was used to obtain two-dimensional crystals of the Botulinum neurotoxin complex, Serotype A and a two-dimensional projection map was reported[12]; however, because the height of the densities was not determined, the masses of the density "blobs" were unknown. In this paper, an improved protein purification and optimized crystallization conditions are reported for the Serotype A botulinum neurotoxin complex. The isolated complexes and smaller components were set up in two-dimensional crystallization trials to observe which components were necessary for crystal formation. Additionally, low-dose electron microscopy on these improved crystals was used to obtain a three-dimensional reconstruction of the botulinum neurotoxin complex. Finally, combining the size and shape of observed densities with previous biochemical observations, we suggest the location of the neurotoxin, non-toxic non-hemagglutinin and hemagglutinin in the complex.

Results and Discussion

Crystallization of individual complex components from the alternative purification

In the crystallization trials of the three different sized complexes of Botulinum neurotoxin complex, only the fractions containing the 500 kD and 900 kD complexes, showed the "triangular" type crystal described previously. In addition to the triangular type crystals, many "single molecules" could be seen adsorbed to the lipid monolayer (Figure 1). Interestingly, these triangular type monomers, which appear to make up the repeating units of triangular crystal, were also seen with protein from the peak containing 500 kD and 900 kD protein. From this observation, the monomer units seen in the crystals probably represent either 500 kD or 900 kD proteins. The crystals formed using the alternate method lacked the necessary quality for a three-dimensional reconstruction. Other eluted components of the botulinum neurotoxin complex obtained from the alternate purification were also set up under various conditions for two-dimensional crystallization; however, none of them have crystallized yet.

Generating the three-dimensional reconstruction of the complex

Using the improved crystallization method, two-dimensional crystals obtained were used to generate the three-dimensional reconstruction. An image of a crystal at 0° tilt is shown in Figure 2. An image of a crystal tilted at +60° is shown in Figure 3. Figure 4 shows in image transform of a crystal tilted at 45°. A three-dimensional density map was produced in both P1 and P3 symmetries. Individually generated images were tested in three different orientations (by rotation of the crystal image by 120 degrees), and the orientation giving the lowest phase residuals were chosen for merging into the final data set. Crystal tilt, the program for selecting the best orientation was written by Dr. Kenneth Downing at Lawrence Berkeley Lab. Structure factors from individual images were used in the final map only if their signal to background ratio was better that 2.5 to 1. The final data set consisted of 109 images with an overall phase residual of 22.602 +/- 8.566 merged in P1, and 24.711 +/- 10.655 merged in P3. Table 1 shows the phase residual of images averaged within 10° tilt increments. Both a P1 and a P3 structure are presented because the actual symmetry awaits further determination.

The average lattice parameters determined for the untilted images are $a^* = 1/136 \pm 1/2$ Å^{-1}, $b^* = 1/139 \pm 1/4$ Å^{-1}, and $\gamma^* = 60° \pm 2°$. Contours used to display the figures of the three-dimensional reconstructions are chosen to envelop approximately a 500 kD protein, assuming that each unit cell contains one copy of the large, hollow triangle and one copy of the smaller globular density on the side. This was calculated assuming the average protein density[13] to be 1.3 g/cm^3. Because the sequence connectivity between the densities is unknown, the boundaries of the unit cell have not been clearly defined. The resolution limit was set at 22 Å during image processing.

The P1 structure has unit cell dimensions of a = 157(±2) Å, b = 159(±4) Å and $\gamma = 120°$ and is shown in Figure 5. A large, ring shaped, central density is seen in the center of Figure 5A surrounded by three smaller globular densities. The length of one edge of the hollow triangle is approximately 112 Å while the length of one edge of the smaller surrounding densities is 67 Å. Figure 5B shows the side view of one of the large hollow triangles and one smaller surrounding density. The longest height of the large ring-like density is 90 Å and the longest height of the small surrounding density is 78 Å. In Figure 5A, asymmetry is seen for both of the densities. The smaller surrounding density actually appears to be a five-pointed star skewed along one side. On the larger hollow triangle, it is seen that the three walls surrounding the hollow center are not equivalent. One wall of the triangle appears to be a somewhat continuous structure of three densities. In Figure 5B, the heights and shape of the three walls also are not equivalent. Two of the three walls have long narrow densities that arch towards the center of the hollow area. Figure 6 shows the same data described above with P3 symmetry imposed; the unit cell parameters are a = b = 157 Å and $\gamma = 120°$.

Calculating the mass of the neurotoxin densities

Further analysis was performed to determine the mass of the neurotoxin densities in the three-dimensional reconstruction. Densitometry on Coomassie stained SDS-PAGE gels of the complex was performed to determine the percent of the complex which was attributed to 150 kD botulinum neurotoxin. The 150 kD botulinum neurotoxin bands are in bands 3 and 4. (The 50 kD band of the neurotoxin runs above the 52 kD band of the multi-subunit

hemagglutinin.) The results of densitometry (Table 2) showed that the botulinum neurotoxin was 28.30% of the whole 900 kD complex. This corresponds to a mass of approximately 260 kD, which would be almost two copies of the 150 kD botulinum neurotoxin per complex molecule. If the mass represented in the 3-dimensional reconstruction shown in Figure 2B is 500 kD, then the 900 kD complex must be made of more than one copy of each density.

Additional support that the mass in the three-dimensional reconstruction represents 500 kD comes from preliminary two-dimensional crystallization of the Serotype B complex. Unlike Serotype A, the Serotype B complex is not secreted in the 900 kD form but exists in 300 kD and 500 kD sizes. Two-dimensional crystals of the Serotype B complex were obtained with approximate unit cell dimensions of a=177.3Å, b=173.2Å and γ=59.6° (Julie Zhang, unpublished data). The triangular crystals of Serotype B complex are seen in Figure 7. Because the Serotype B crystals (Figure 7), containing a mixture of 300 kD and 500 kD species, share the same triangular shape as the Serotype A crystals (Figure 2), containing a mixture of 500 kD and 900 kD species, the crystallized monomer is likely the 500 kD species.

Our results indicating that the 900 kD complex may be a dimer of the 500 kD complex comport with results from the literature. Under sucrose density gradient centrifugation at pH 6.0 and in 1000 mM sodium chloride, the 500 kD and 900 kD complexes sediment to the same position corresponding to 500 kD[7]. Furthermore, the 900 kD protein runs identically to the 500 kD protein on SDS-PAGE, except that the stoichiometry of the 35 kD protein is increased [14]. The increased stoichiometry implies that the 900 kD complex may be a dimer of the 500 kD protein linked via extra copies of the 35 kD protein.

Such a "building-block" approach to the construction of the 900 kD complex may be seen in the formation of the 500 kD and 300 kD complexes, which appears to be controlled by the 120 kD non-toxic, non-hemagglutinating protein. Although the function of the 120 kD non-toxic non-hemagglutinin protein is uncertain, its status as nicked or unnicked determines the size of the complex[14]. The 120 kD protein may be post-translationally nicked into 106 kD and 14 kD fragments [15]. The nicked fragments appear to associate via noncovalent interactions [14]. The 300 kD complex contains one 150 kD neurotoxin molecule and one *nicked* 120 kD non-toxic non-hemagglutinin molecule . Interestingly, the 500 kD

complex is formed by the association of an *unnicked* 120 kD non-toxic non-hemagglutinin, a 150 kD neurotoxin and additional hemagglutinating proteins. From these observations, which also hold for Serotype C and D toxins, the post-translational nicking of the 120 kD band may prevent the hemagglutinating protein from binding to the 300 kD toxin [14]. Recent evidence indicates that the 35 kD hemagglutinin not only provides the principal source of hemagglutination ability but also provides protection against proteases found in the gastrointestinal tract[16].

Identifying the densities in the unit cell

Using the biochemical data above and the three-dimensional structure of the 150 kD botulinum neurotoxin, the identities of densities in the three-dimensional reconstruction are proposed. The three-dimensional structure of 150 kD neurotoxin was determined by x-ray crystallography from three-dimensional crystals set up at pH 7[17]. The three-dimensional reconstruction was determined by electron microscopy using 2-dimensional crystals at pH 5. The differences between these two pH's may be structurally significant since the shape of the translocation domain may vary as a function of pH. Going from pH 7 to 5, the translocation domain supposedly forms a pore across the endosomal membrane[18], allowing the catalytic domain to escape into the cytosol. Thus, the structure of the 150 kD botulinum neurotoxin at pH 7 would probably not fit exactly into a map generated at pH 5.

The three-dimensional structure of 150 kD botulinum neurotoxin fits approximately into two regions of the three-dimensional map. The 150 kD botulinum neurotoxin structure appears as three domains arranged linearly[17]. The longest length of the structure is 115 Å. The longest height corresponds to the translocation domain, a long narrow structure of height 100 Å. There are two locations in the central hollow triangle in which the 150 kD botulinum neurotoxin may fit. The first location is in the long continuous wall of the triangle shown in figure 8A and 8B. Figure 8B shows that the translocation domains of the two structures are in different conformations. The pH 5 map shows a bend in the translocation domain to one side and towards the center of the hollow core.

The second location is shown in Figure 9A and 9B. Again, the translocation domain does not fit perfectly into the density in Figure 9B. The long helices in the x-ray structure of the translocation domain would fit into the long narrow density seen in the electron microscopy structure if one pictures a sharp bend inward on the ends of the translocation domain. To determine whether the 150 kD botulinum neurotoxin fits into the first or second location, an antibody-bound structure or a higher resolution density map of the botulinum neurotoxin is required.

As discussed above, the 150 kD botulinum neurotoxin and the 120 kD non-toxic non-hemagglutinin make up the 300 kD complex. Only when the 120 kD non-toxic non-hemagglutinin is not nicked into 106 kD and 14 kD bands, does the multi-subunit hemagglutinin bind and form the 500 kD and 900 kD complexes[15]. These observations may suggest that the 150 kD neurotoxin is directly associated with the non-toxic non-hemagglutinin while the multi-subunit hemagglutinin associates with the unnicked 120 kD non-toxic non-hemagglutinin.

Additional support for the interaction between the 120 kD non-toxic non-hemagglutinin comes from observations of botulinum neurotoxin chromatography. The three components of the botulinum neurotoxin may be separated by ion exchange chromatography[19]. At salt concentrations above 150 mM and pH above 7.4, the 150 kD botulinum neurotoxin is separated from the non-toxic proteins. At higher salt concentrations, the non-toxic proteins, composed of the 120 kD non-toxic non-hemagglutinin and the multi-subunit hemagglutinin, elute as one peak and remain complexed together. The affinity between the non-toxic proteins suggests that there is an interaction between these two components.

Proposed Model Identifying Densities in 3-Dimensional Reconstruction

Based on the biochemical evidence, a model identifying the densities of the three-dimensional reconstruction is proposed (Figure 10). As stated in the previous section, because the botulinum neurotoxin molecule appears to associate closely with the 120 kD non-toxic non-hemagglutinin in the 300 kD complex, the density of the central hollow triangle is likely

the 150 kD botulinum neurotoxin surrounded by the 120 kD non-toxic protein. The location in Figure 8 is preferred because it allows the 120 kD protein to have maximum interaction with the globular density on the side, which must be the multi-subunit hemagglutinin. The presence of the 120 kD non-toxic non-hemagglutinating protein and the hemagglutinin would protect the vulnerable neurotoxin from proteolysis by providing a "protein armor." The images shown in Figures 8 and 9 depict the 500 kD complex outlined in yellow. Since the 900 kD complex contains two copies of 150 kD botulinum neurotoxin, it would consist of two large hollow triangles, composed of the 150 kD neurotoxin and the 120 kD non-toxic non-hemagglutinin complexed via an unknown number of multi-subunit hemagglutinins. This model of the 150 kD neurotoxin nestled into the 120 kD non-toxic non-hemagglutinin refines the crude antibody map developed in previous work [20] and provides a framework to explain the neurotoxin's amazing ability to withstand the furious defenses of the gastrointestinal tract.

Biological Implications

The organism *Clostridium botulinum* produces the 150 kD botulinum neurotoxin in complexes of three sizes: 300 kD, 500 kD and 900 kD. The complexes consist of the neurotoxin, proteins associated with a hemagglutinating ability, and non-toxic non-hemaggluting proteins whose function is unknown. The botulinum neurotoxin complex is one of the most potent oral toxins, attacking the neuromuscular junction and producing paralysis and death. Botulinum neurotoxin's toxicity increases as the mass of the complex increases, and, without any complexed proteins, the 150 kD neurotoxin would be rapidly degraded by the pH and proteases found in the gastrointestinal tract. Little structural information is known about the neurotoxin complex, and we present a three-dimensional reconstruction of the complex visualized using electron microscopy.

The reconstruction shows the electron density of the 500 kD complex. In the reconstruction, the neurotoxin is surrounded by the non-toxic non-hemagglutinating protein and the hemagglutinating protein. The presence of the non-toxic non-hemagglutinating and

hemagglutinating proteins may shield vulnerable regions of the neurotoxin from the harsh pH and proteases found in the gastrointestinal tract.

Materials and Methods

Alternate protein purification of complexes

Following a purification procedure from Inoue et al,[14] the following proteins could be isolated: 35 kD protein, hemagglutinating protein complex (15 kD, 19 kD, 20 kD, 35 kD and 52 kD), 120 kD non-toxic non-hemagglutinating protein, 300 kD complex, 500 kD complex and 900 kD complex. Forty milligrams of crude cell precipitate was obtained from Dr. Bibhuti DasGupta. The crude cell precipitate was dialyzed into 50 mM sodium acetate, pH 4.2 and loaded on SP-Toyopearl 650M cation exchange resin (Tosohaas, Inc., Japan). The column was then washed with the same buffer, and a sodium chloride gradient to 500 mM was applied. Five protein peaks were collected. Peak 1 corresponded to 120 kD non-toxic non-hemagglutinin. Peak 2 contained free 35 kD protein. Peak 3 was the hemagglutinating protein complex (15 kD, 19 kD, 20 kD, 35 kD and 52 kD). Peak 4 corresponded to a mixture of 500 kD and 900 kD complex. Finally, peak 5 was the 300 kD complex. An attempt to crystallize the protein in each peak was performed. Proteins were dialyzed or diluted into pH 5.0, 50 mM sodium acetate buffer with 100 mM sodium chloride and two-dimensional crystals were set up as described above. The specimen for each crystallization attempt was harvested after 48 hours and examined on the Zeiss 10A at 80 keV.

Improved protein preparation and two-dimensional crystallization

Protein preparation was modified from that previously described[12]. The 900 kD complex was obtained as an ammonium sulfate precipitate at 3.3 mg/ml in 50 mM sodium citrate pH 5.5. Before use, the precipitated complex was centrifuged at 26890 × g (Eppendorf Microfuge) for 20 minutes, resuspended and dialyzed against 50 mM sodium acetate (Fluka, Ronkonkoma, NY), 100 mM sodium chloride (Fluka), 1 mM sodium azide, pH 5.0. Concentration was determined by A_{278} measurements (1.66 AU/mg ml^{-1}) using a Shimadzu UV-160 spectrophotometer [21]. Final protein concentration was approximately 1 mg/ml.

Larger, more coherent crystals than previously discussed were obtained using phosphatidyl glycerol (POPG) rather than phosphatidyl choline (POPC). Lyso-lactosyl ceramide (lyso-LC) was used instead of lactosyl ceramide. The difference in these two lipids is that lyso-lactosyl ceramide lacks one of the fatty acid chains, possibly allowing the lipid to diffuse and pack more easily. The ratios of lyso-LC to POPG were varied. Trials consisted of: 100% lyso-LC:0% POPG, 95% lyso-LC:5%POPG, 85% lyso-LC:15% POPG, 75% lyso-LC:25% POPG and 65% lyso-LC:35% POPG. Lipid composition was optimized at 85% lyso-LC:15% POPG at 1 mg/ml. New batches of lipid and protein were made every three to four weeks.

Botulinum neurotoxin complex was diluted into 50 mM sodium acetate, pH 5.0, 100 mM sodium chloride, 1 mM sodium azide to a concentration of 0.08 mg/ml. Next, 50 µl of the diluted protein solution was pipetted into Teflon wells of approximately 1.5 mm deep and 7 mm in diameter. 1 µl of 85% lyso-LC:15% POPG solution was placed onto the protein solution[22]. Specimens were incubated for 48 to 72 hours at room temperature in a sealed humid chamber on an anti-vibration surface. Shortly before harvesting, carbon coated electron microscopy grids were rendered hydrophobic by glow discharge treatment described previously and placed horizontally onto the surface of the protein drop. The grid was removed from the surface of the drop, washed with 50 µl filtered water and stained with 50 µl filtered 1.5% (w/v) uranyl acetate (Fluka, Ronkonkoma, NY) for approximately 60 seconds. The grid was stored in a dark, dry environment until viewing on the microscope.

Data collection and processing for three-dimensional reconstruction

Specimens were examined in a JEOL 4000 at an accelerating voltage of 400 keV, and images were recorded on Kodak SO-163 film. Images were collected under low-dose conditions of approximately 10 electron per $Å^2$ at 50 K magnification. A total of 512 images were collected, of which 1/3 were taken using spot-scan focus ramp imaging to minimize the effects of specimen drift and maintain focus for highly tilted specimens. Of the 512 images taken, approximately 250 were digitized with a Perkin-Elmer PDS 1010M flat bed microdensitometer with a step size of 20 µm in arrays of 1000 by 1000 elements. The tilt angle of the images ranged from –60° to +60°. Each image was corrected for lattice distortion [23] and processed using standard processing programs[24].

References

1. Gill, D.M., *Bacterial toxins: a table of lethal amounts.* Microbiol Rev, 1982. **46**: p. 86-94.
2. Chen, F., G.M. Kuziemko, and R.C. Stevens, *Biophysical characterization of the stability of the 150-kilodalton botulinum toxin, the nontoxic component, and the 900-kilodalton botulinum toxin complex species.* Infection and Immunity, 1998. **66**(6): p. 2420-2425.
3. Sugii, S., I. Ohishi, and G. Sakaguchi, *Correlation between oral toxicity and in vitro stability of Clostridium botulinum type A and B toxins of different molecular sizes.* Infection and Immunity, 1977. **16**(3): p. 910-4.
4. Singh, B.R., B. Li, and D. Read, *Botulinum versus tetanus neurotoxins: Why is botulinum neurotoxin but not tetanus neurotoxin a food poison?* Toxicon, 1995. **33**(12): p. 1541-1547.
5. Ohishi, I., S. Sugii, and G. Sakaguchi, *Oral toxicities of Clostridium botulinum toxins in response to molecular size.* Infection and Immunity, 1977. **16**(1): p. 107-9.
6. Sakaguchi, G., I. Ohishi, and S. Kozaki, *Botulism- Structure and Chemistry of Botulinum*, in *Bacterial Toxins*, M.C. Haidegree and A.T. Tu, Editors. 1988, Marcel Dekker, Inc.: New York. p. 191-208.
7. Sugii, S. and G. Sakaguchi, *Molecular construction of Clostridium botulinum type A toxins.* Infection and Immunity, 1975. **12**(6): p. 1262-70.
8. Lamanna, C., *Hemagglutination by Botulinal Toxin.* Proc. Soc. Exp. Biol. Med., 1948. **69**: p. 332-336.
9. Boroff, D.A., S. Nyberg, and S. Hoglund, *Electron microscopy of the toxin and hemagglutinin of type A Clostridium botulinum,.* Infect. Immun., 1972. **6**(6): p. 1003-1007.
10. Morgan, D.G., *et al.*, *Two dimensional crystals of botulinum toxin, serotype B,.* Proc. of the Ann. Meeting of the Electron Microscopy Society of America, 1989: p. 1034-1035.
11. Schmid, M.F., J.P. Robinson, and B.R. DasGupta, *Direct visualization of botulinum neurotoxin-induced channels in phospholipid vesicles.* Nature, 1993. **364**: p. 827-830.
12. Burkard, F., *et al.*, *Electron Density Projection Map of the Botulinum Neurotoxin 900-Kilodalton Complex by Electron Crystallography.* Journal of Structural Biology, 1997. **120**: p. 78-84.
13. Cantor, C.R. and P.R. Schimmel, *Part 1. The conformation of biological macromolecules*, in *Biophysical Chemistry*. 1980, W.H. Freeman and Company: New York, NY. p. 115-116.
14. Inoue, K., *et al.*, *Molecular composition of Clostridium botulinum type A progenitor toxins.* Infection and Immunity, 1996. **64**(5): p. 1589-1594.
15. Fujita, R., *et al.*, *Molecular characterization of two forms of nontoxic-nonhemagglutinin components of Clostridium botulinum type A progenitor toxins.* FEBS Letters, 1995. **376**(1-2): p. 41-44.
16. Fu, F.N., S.K. Sharma, and B.R. Singh, *A protease-resistant novel hemagglutinin purified from type A Clostridium botulinum.* Journal of Protein Chemistry, 1998. **17**(1): p. 53-60.
17. Lacy, D.B., *et al.*, *Crystal Structure of Botulinum Neurotoxin Type A and Implications for Toxicity.* Nature Structural Biology, 1998. **In press**.
18. Blauestein, R.O., *et al.*, *The N-terminal half of the heavy chain of botulinum neurotoxin forms channels in planar phospholipid bilayers.* FEBS. Lett., 1987. **226**: p. 115-120.
19. Dasgupta, B.R. and D.A. Boroff, *Separation of Toxin and Hemagglutinin from Crystalline toxin of Clostridium botulinum Type A by Anion Exchange Chromatography and Determination of Their Dimensions by Gel Filtration.* J. Biol. Chem., 1968. **243**: p. 1065-1072.
20. Chen, F., *et al.*, *Antibody mapping to domains of botulinum neurotoxin serotype A in the complexed and uncomplexed forms.* Infection and Immunity, 1997. **65**(5): p. 1626-1630.

21. Knox, J.N., W.P. Brown, and L. Spero., *The role of sulfhydryl groups in the activity of type A botulinum toxin.* Biochim. Biophys. Acta., 1970. **214**: p. 350-354.
22. Darst, S.A., *et al., Two-dimensional crystals of Escherichia coli RNA polymerase holoenzyme on positively charged lipid layers,.* J. Mol. Biol., 1988. **203**: p. 269-273.
23. Henderson, R., *et al., Structure of purple membrane from Halobacterium halobium: Recording, measurement and evaluation of electron micrographs at 3.5 a resolution,.* Ultramicrsocopy, 1986. **19**: p. 147-178.
24. Crowther, R.A., R. Henderson, and J.M. Smith, *MRC Image processing programs,.* J. Struct. Biol., 1996. **116**: p. 9-16.

Tables for the 3D-EM Paper

Table 1 Phase Residuals of Images Included in 3-Dimensional Map averaged in 10° increments					
Tilt Angle Range	Number of Images Averaged	P1 Symmetry		P3 Symmetry	
		Average Phase Residual	Standard Deviation	Average Phase Residual	Standard Deviation
0°–9°	21	15.040	6.313	15.180	5.948
10°–19°	18	18.199	5.485	18.174	6.182
20°-29°	8	16.617	3.262	15.969	4.428
30°–39°	17	24.387	4.486	25.879	5.093
40°–49°	28	26.516	6.501	31.179	7.946
50°-60°	17	30.891	9.920	35.524	11.906

Table 2. Densitometry of the 900 kD Botulinum Neurotoxin Complex

Band Number	Molecular Weight (kD)	Intensity (Optical Density)	Relative Percent (%)
1	120	0.361	11.37
2	106	0.060	1.89
3	100	0.553	17.39
4	55	0.347	10.91
5	52	0.717	22.55
6	35	0.973	30.63
7	19, 21	0.095	2.98
8	15	0.072	2.28

Figure 1

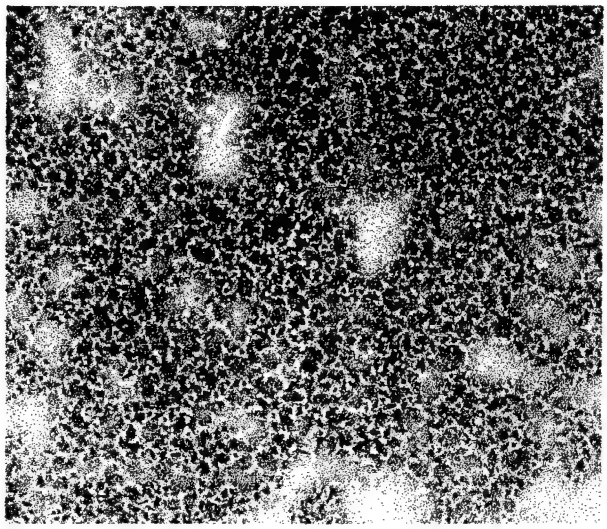


Figure 2

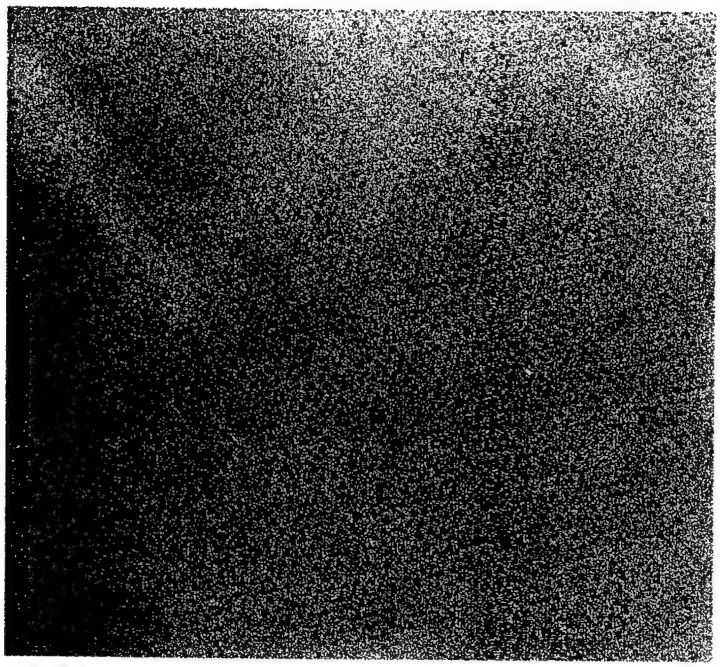


Figure 3

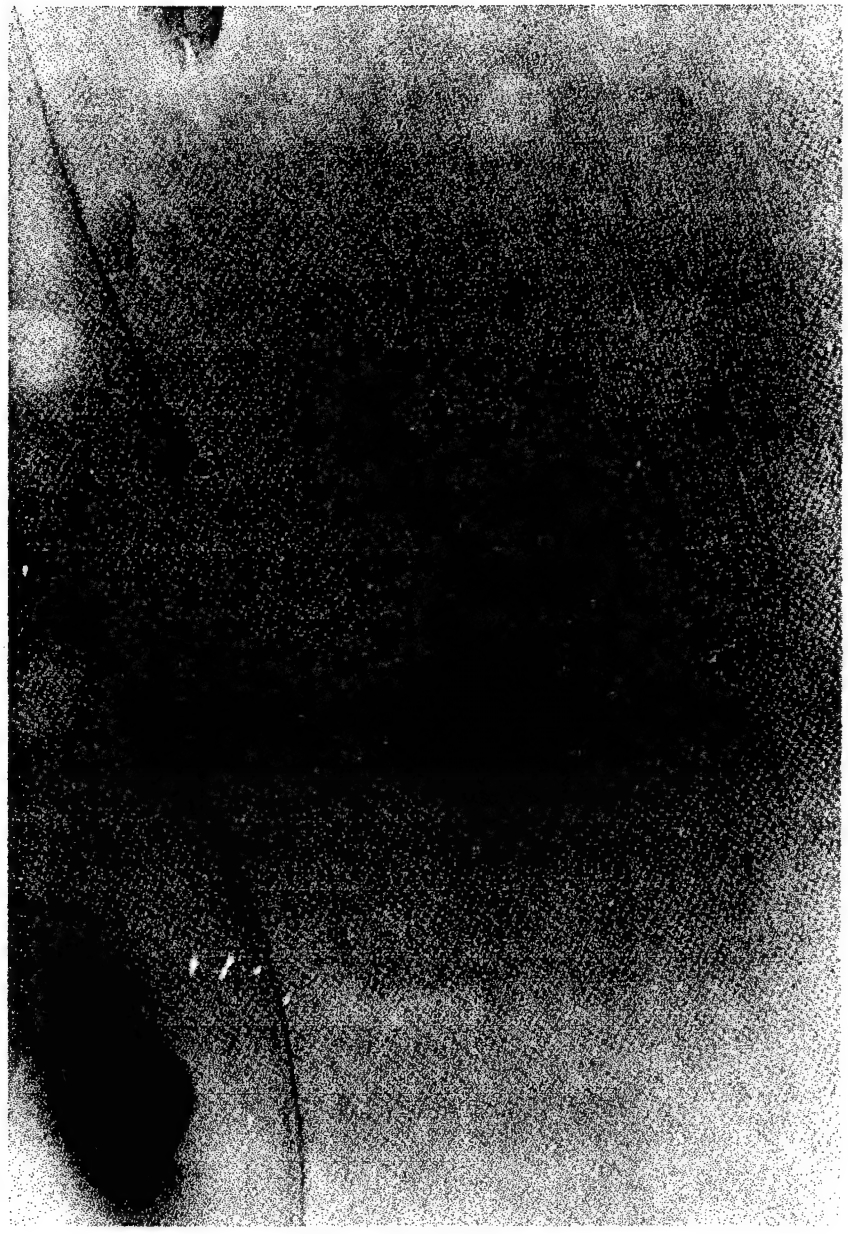


Figure 4

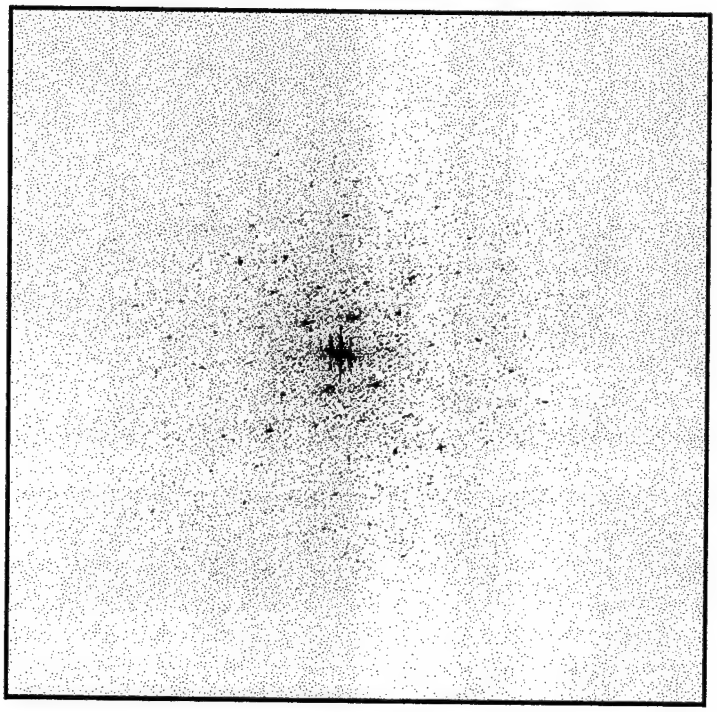
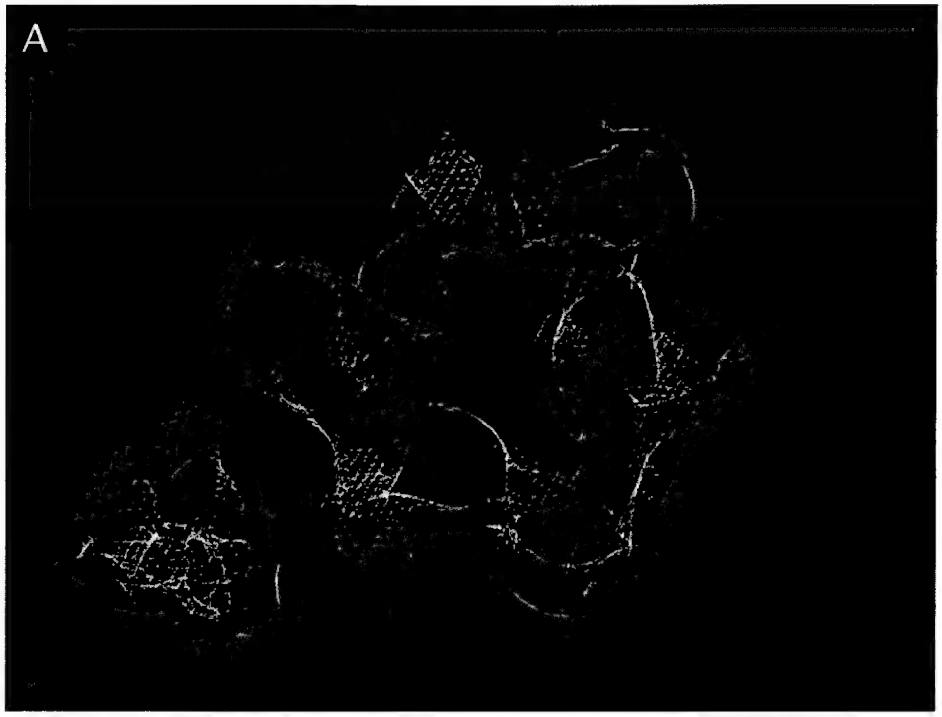


Figure 5



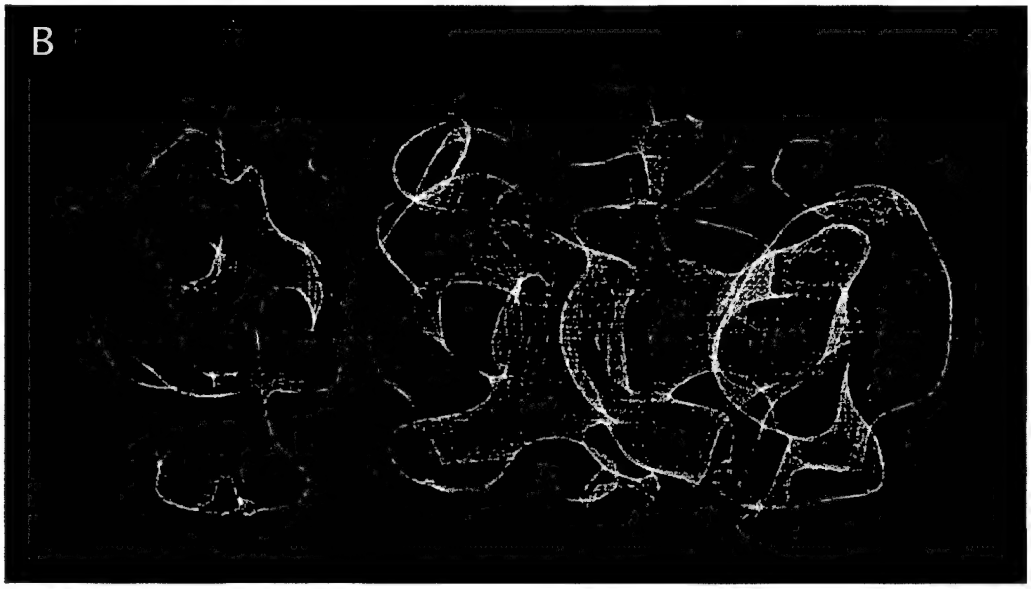


Figure 6

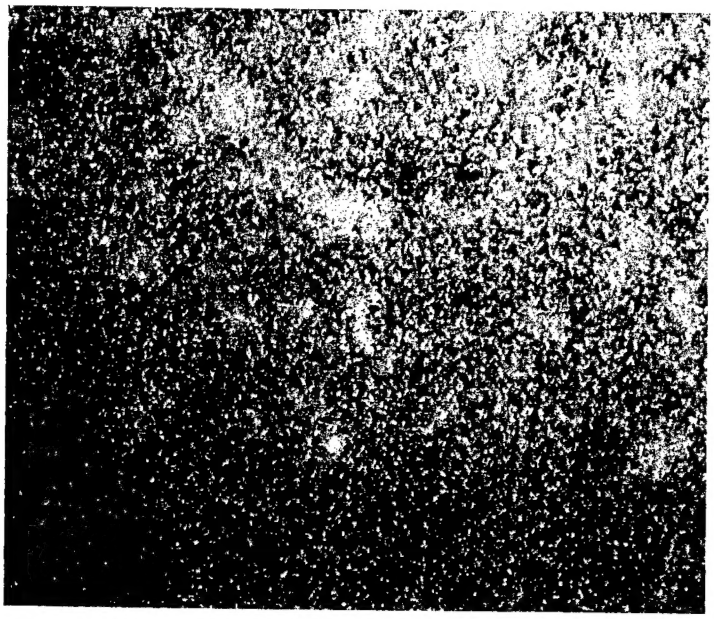
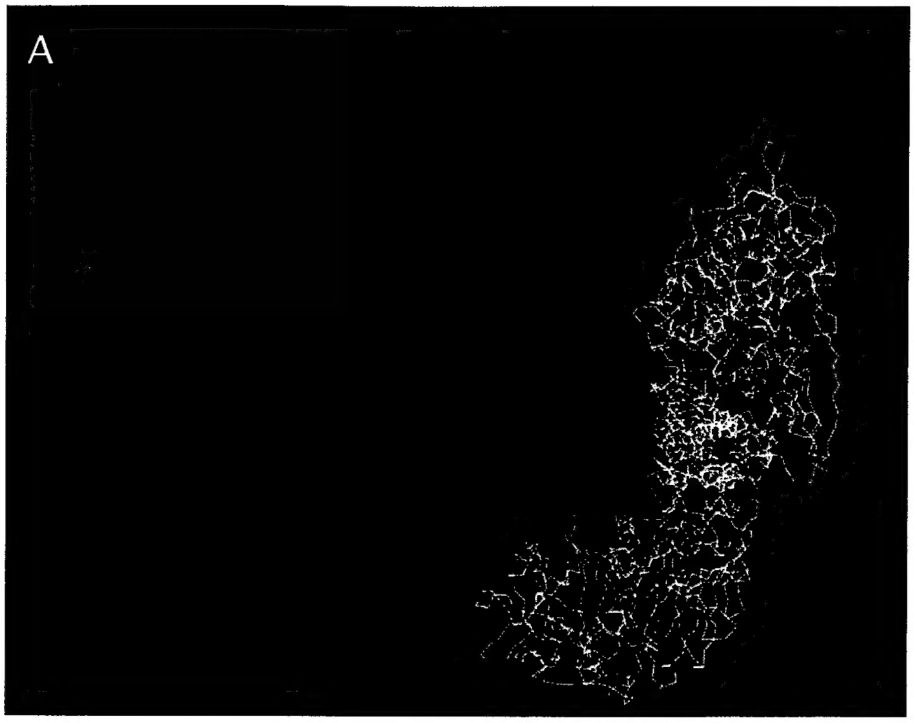


Figure 7



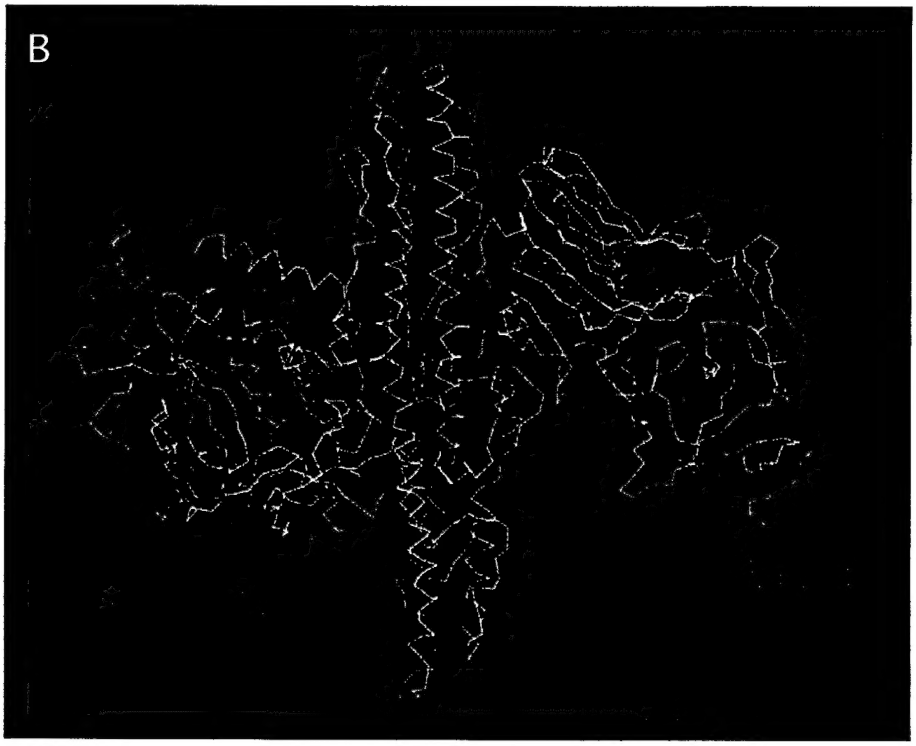


Fig 9

Figure 8

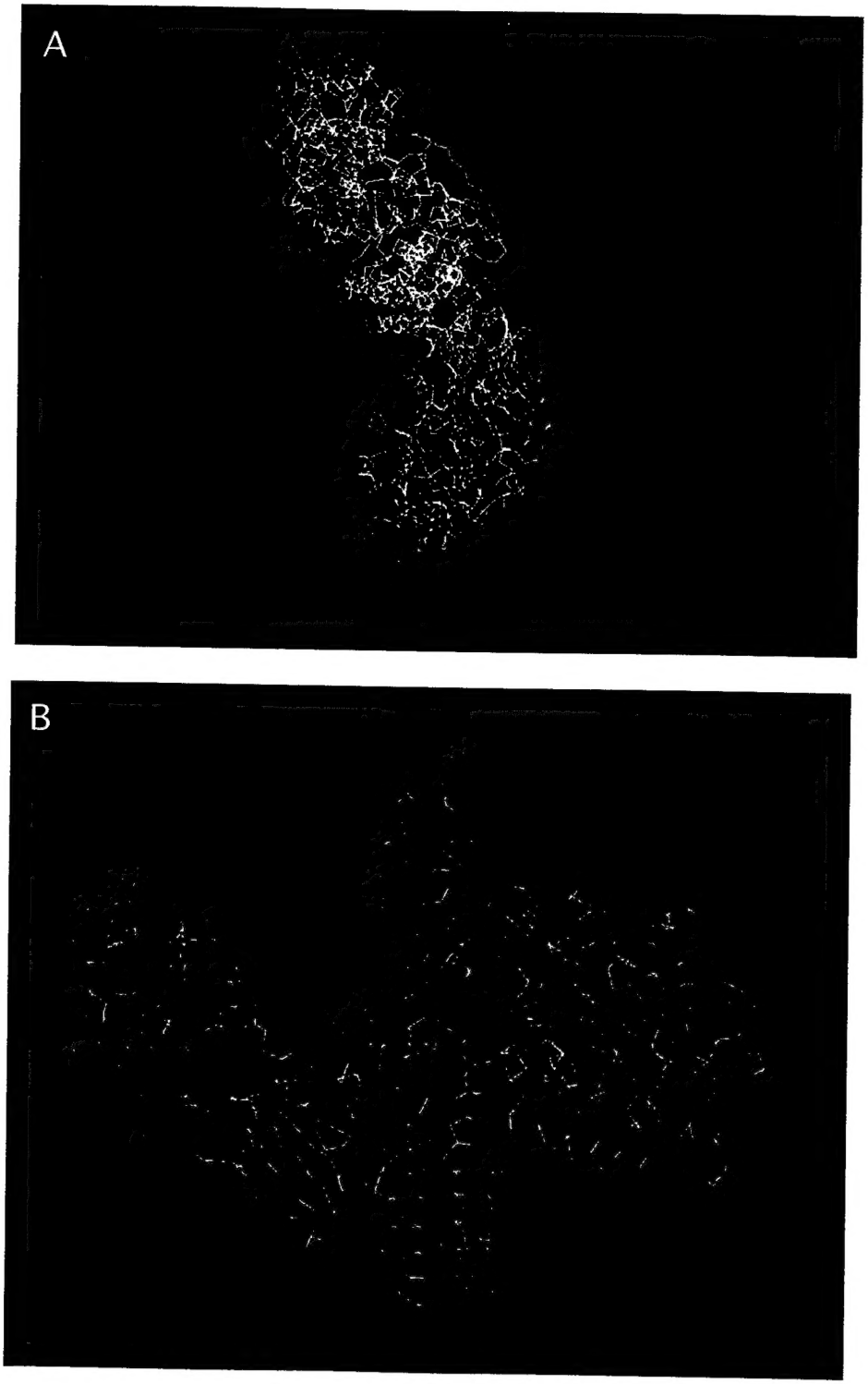


Figure 9